UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS FÍSICAS

TESIS DOCTORAL

Land–air interactions and subsurface heat transport: the role of external forcing and implications for last millennium temperature reconstructions

Interacciones tierra–aire y el transporte de calor en el subsuelo: el papel de los forzamientos externos e implicaciones para la reconstrucción de temperaturas del último milenio

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Camilo Andrés Melo Aguilar

DIRECTOR

Jesús Fidel González Rouco

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS FÍSICAS

TESIS DOCTORAL

Land–air interactions and subsurface heat transport: the role of external forcing and implications for last millennium temperature reconstructions

Interacciones tierra–aire y el transporte de calor en el subsuelo: el papel de los forzamientos externos e implicaciones para la reconstrucción de temperaturas del último milenio

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Camilo Andrés Melo Aguilar

DIRECTOR

Jesús Fidel González Rouco

Land–air interactions and subsurface heat transport: the role of external forcing and implications for last millennium temperature reconstructions

Interacciones tierra–aire y el transporte de calor en el subsuelo: el papel de los forzamientos externos e implicaciones para la reconstrucción de temperaturas del último milenio

Memoria que presenta
Camilo Andrés Melo Aguilar
para optar al grado de
Doctor en Ciencias Físicas

Director:
Dr. J. Fidel González Rouco

Departamento de Física de la Tierra y Astrofísica
Facultad de Ciencias Físicas
Universidad Complutense de Madrid

A Rebeca y Guillermo

Agradecimientos

Han pasado casi 5 años desde el inicio de esta ardua y tortuosa tarea en la cual me embarqué casi por causa del azar. En mi vuelta a Madrid en el verano de 2015, se me presentó la oportunidad de iniciar esta investigación sin haber estado buscándola. A pesar del gran reto que significaba para mi, decidí emprender este proceso, motivado por el mismo reto que este significaba. Ha sido este un camino de altibajos emocionales, pero sobre todo de gran aprendizaje. Es por esto que quiero reservar unas pocas líneas para mostrar mi agradecimiento a las personas que han hecho posible culminar esta labor y con las cuales he compartido este tiempo.

En primer lugar quiero agradecer a Rebeca porque sin su apoyo es posible que hubiera claudicado ya hace algún tiempo. En momentos de gran dificultad personal, por el hecho de encontrarme tan lejos de mi gente y de mis costumbres, ella me ha soportado (en todos los sentidos) y me ha dado motivos para continuar. Es por ella por quien hoy es posible terminar esta tarea.

Durante este camino por la universidad he tenido la oportunidad y el gusto de conocer a Fidel, que ha sido un ejemplo para mi no sólo como director de este proyecto, sino más que todo como persona. Su dedicación y amor por el trabajo es inigualable. Gracias por haberme enseñado tantas cosas y por la paciencia que has tenido conmigo. Finalmente hemos logrado sacar esta investigación adelante que se encontraba guardada en el cajón durante varios años esperando por mi. Al final ha valido la pena. Quiero darte las gracias porque en algunas ocasiones has tirado

de mi poniéndote el traje de “jefe” que te corresponde, pero en otras también el de amigo. Gracias por las ocasiones en que compartimos una copa discutiendo del trabajo y de la vida. A pesar de que tu tiempo era escaso, siempre hubo tiempo para esto, y donde sea que estuviéramos, encontramos alguna trampa para turistas donde refugiarnos.

A la gente del grupo Palma, con quienes he compartido tiempo y espacio, quisiera también expresar mi gratitud por haberlos conocido. Lo haré por orden cronológico. Jorge Navarro, tu buen sentido del humor siempre hizo agradable la hora de la comida. Cuando la conversación se empantanaba en temas de política, tus chistes me alegraban el rato. Marisa, ha sido un lujo compartir este tiempo contigo. Elena, tu buena energía se sentía siempre en el ambiente. Jorge Álvarez, se sentía la pasión por tu trabajo, eso es de admirar. Alex, siempre de buen rollo, un gusto conocerte. Etor, eres único, sabes de todo, me alegra haberte conocido. Espero que esta nueva etapa te llene de alegrías. Cristina V., siempre tenías una sonrisa a mano, espero que el confinamiento haya servido por lo menos para acabar con las series de Neflix. Ilaria, super maja y super lista, me has sacado 2 años de ventaja, ojalá que nos veamos de nuevo para tomar alguna cerveza en Alemania, o en Italia. Norman, tu llegada al grupo ha sido una maravilla, nuestras reuniones semanales han contribuido con mi trabajo, además he disfrutado los vermuts que nos tomamos por Madrid. Ánimo que eres el siguiente. Cristina R., un espíritu alegre, me encanta que sientas tanta curiosidad por Colombia. Espero que algún día puedas ir y ojalá este yo allí para enseñarte algunas cosas guapas al estilo colombiano. Javi, no pudimos celebrar nuestras tesis como es debido, en cuanto haya mas tranquilad nos vemos en las Fallas para festejar un poco, o si puede ser, también en Colombia, nunca se sabe. Deseo que vengan cosas chulas para ti en el futuro. La nueva generación viene con fuerza. Felix, muchos temas interesantes por delante, seguro que lo harás muy bien, ánimo. Dani, disfruta de este tiempo que se pasa volando. Aparte de la gente del grupo, también le quiero agradecer a Johann por haberme recibido en Hamburgo esos 3 meses y por las caminatas por el río, he disfrutado el tiempo allí.

Finalmente dedicar unas palabras a mi familia, los pienso mucho desde la distancia, me hacen mucha falta. Gracias por haberme enseñado que hay que hacer siempre lo que a uno le de la gana.

Acknowledgements

Camilo Melo Aguilar was supported by an FPI grant (formación de personal investigador) no. BES-2015-075019 from MINECO (Ministerio de Económia Industria y Competitividad), and projects IlModels CGL2014-59644-R and GreatModelS RTI2018-102305-B-C21.
A short stay of 3 months at the Max Planck Institute for Meteorology (MPI, Germany) was funded by the FPI grant no. BES-2015-075019. Thanks to the MPI and particularly to Dr. J. Jungclaus who made it possible, and also to Dr. S. Lorenz for their scientific and technical support.
The model output employed for the development of the research papers Melo-Aguilar et al. (2018) and Melo-Aguilar et al. (2020) that are contained in Chapters 4 and 5 was obtained from the Community Earth System Model Last Millennium Ensemble (CESM–LME; Otto-Bliesner et al., 2016) developed within the CESM1(CAM5) Last Millennium Ensemble Community Project and supercomputing resources provided by NSF/CISL/Yellowstone.
The ERA5-Land reanalysis used in Chapter 3 is funded under the Copernicus Regulation and operated by ECMWF under the ECMWF Agreement.
The soil temperature observations for the Iberian Peninsula and Balearic Islands that were used in the analyses developed in Chapter 3 were provided by the Spanish Meteorological Agency (AEMET).

Contents

Agradecimientos V

Summary / Resumen XI

List of acronyms XXIII

1 Introduction 1
- 1.1 Land–atmosphere coupling 1
- 1.2 The subsurface thermal regime 2
- 1.3 Surface processes influencing the land-air temperature coupling 6
  - 1.3.1 Model simulations to assess the long-term land–air temperature coupling 9
- 1.4 Borehole-based temperature reconstructions 12
- 1.5 Assessing borehole temperature reconstructions 15
- 1.6 Main objectives and structure of the thesis 19

2 Data: observations and model simulations 23
- 2.1 Soil temperature observational dataset for Spain (SoTOS) 23
  - 2.1.1 Quality control 27
  - 2.1.2 Final quality controlled daily surface air and soil temperature dataset 32
- 2.2 ERA5-Land reanalysis 34
- 2.3 Community Earth System Model–Last Millennium Ensemble (CESM-LME) 35
- 2.4 Conclusions 38

3 Shallow soil thermal regime and SAT–GST coupling: a case study over Spain ... 41
3.1 Methods ... 42
3.2 Results ... 45
3.2.1 SAT and ST climatologies ... 45
3.2.2 Subsurface thermal regime ... 50
3.2.3 SAT-GST coupling ... 60
3.3 Conclusions ... 70

4 Influence of radiative forcing factors on the SAT–GST relationship during the LM ... 73
4.1 Methods ... 74
4.2 Results ... 76
4.2.1 SAT–GST long-term changes ... 88
4.3 Conclusions ... 103

5 Methodological and physical biases in borehole temperature reconstructions ... 107
5.1 Experimental design ... 108
5.1.1 Forward and inverse models ... 109
5.1.2 Pseudo-proxy set up ... 110
5.2 Results ... 113
5.2.1 Methodological issues ... 114
5.2.2 Biases of past SAT borehole-based reconstructions due to physical processes ... 123
5.3 Conclusions ... 132

6 Conclusions and outlook ... 137
6.1 Main conclusions ... 137
6.2 *Outlook* ... 141

References ... 143

Land–air interactions and subsurface heat transport: the role of external forcings and implications for last millennium temperature reconstructions

The energy exchanges at the land-atmosphere interface play a fundamental role in several weather and climatic processes both in the atmosphere and below the land surface. A part of the radiation that reaches the surface is reflected away or dissipated by sensible and latent heat fluxes back to the atmosphere while the ground heat flux balances the excess of incoming energy. This surface energy balance establishes a link between the atmosphere and the ground surface that often translates into a strong coupling between the surface air temperature (SAT) and the ground surface temperature (GST). In turn, the surface temperature stands as the key variable since it affects the amount of heat flowing into the soil through thermal conduction, thus controlling the distribution of temperature below the surface. The coupling between SAT and GST and the conductive heat transport of the surface temperature into the soil stand as the central assumptions for reconstructing the past surface temperature variations from present-day subsurface temperature measurements in a technique known as borehole reconstructions. As with every type of temperature reconstruction method, the borehole technique is subject to uncertainty sources principally due to the influence of physical processes at the surface that interrupt the SAT–GST coupling assumption over a diversity of temporal scales. Understanding the physical mechanisms behind such processes is crucial to gain confidence in the borehole technique estimates regarding the past SAT variations.

Objectives of the study

The main goal of this Thesis is to improve our understanding of the interactions between the ground surface and the atmosphere that influence the coupling between the surface air temperature (SAT) and ground surface temperature (GST) on various temporal and spatial scales. Also, this works aims at evaluating the characteristics of the subsurface thermal regime. The assumptions that the air-ground coupling is strong and stable and that the subsurface thermal regime is conductive are the central hypotheses that allow recovering past surface temperature changes from borehole temperature profiles (BTPs). To address these issues, observational and model data have been used. A surface air and soil temperature observational database for Spain has been developed. The development of a quality controlled dataset of soil temperatures is, in itself, a novel contribution given the scarcity of such products. This dataset allows exploring the subsurface heat transport and the SAT–GST coupling for different soil types and climatological conditions on interannual and decadal timescales. Furthermore, a recent ensemble of last millennium (LM) simulations with the Community Earth System Model (CESM) is employed to explore further the physical mechanisms that control the energy fluxes at the ground-atmosphere interface on multi-centennial timescales and different spatial scales. Likewise, using this ensemble of simulations also helps understand the individual effects of the external forcings on the land-atmosphere interactions and their implications on the SAT–GST coupling. To our knowledge, this has not been addressed earlier in long millennial simulations in which the effect of each external forcing is considered.

Finally, this work aims at evaluating the overall performance of the borehole temperature reconstruction technique to recover long-term surface air temperature variations over the LM in a pseudo-proxy approach. This is focused on quantifying the effects of a set of methodological constraints that may distort the recovery of past GSTs from BTPs, including the spatial and temporal of available BTPs. Further, this work explores the physical influences on SAT–GST long-term decoupling on retrieving the pas SAT variations from subsurface temperature measurements. The latter has not been systematically addressed before.

Main results and conclusions

A dataset of 39 sites monitoring soil temperatures at levels 0.05, 0.1, 0.2, 0.5 and 1 m depth has been compiled and quality controlled and will constitute a useful contribution for the community of scientists interested in soil temperatures

and land-air interactions. The analysis of this dataset has shown that heat transport is mainly dominated by conduction within the shallow subsurface despite non-conductive processes that also influence the subsurface thermal regime. The analysis of the connection between SAT and GST using observational data indicates that, generally, there is a strong coupling over peninsular Spain. However, due to the effect of seasonal processes, as the radiative heating of the ground surface in the summer months in combination with vegetation and moisture effects, there is a large offset between SAT and GST that alter the strong coupling assumption. Further, it has been shown that at decadal timescales, SAT–GST decoupling may develop over central Spain due to a larger temperature increase of GST relative to SAT during the last few decades. The main reason for such an SAT–GST decoupling may be found in decreasing precipitation trends and surface humidity that limit the available soil moisture content, especially in the summer months. As less water is available at the soil surface to evaporate, the high rates of incoming solar radiation heat the ground surface and lead to an overall response of larger GST warming relative to the SAT.

An important aspect of this Thesis is to evaluate if the SAT–GST coupling can also be corrupted at longer timescales. The CESM-LME simulations showed a strong SAT–GST coupling at a global scale and above multi-decadal timescales since GST closely tracks SAT throughout the LM. Nevertheless, at regional scales, there are several situations in which long-term SAT–GST decoupling may exist due to the effects of external forcings. This includes modifications of the energy fluxes at the ground-air interface due to land use land cover changes. Similarly, long-term snow cover feedbacks due to the interaction of greenhouse gases warming with other external forcings can lead to long-term SAT–GST decoupling during the industrial period. These local to regional SAT-GST decoupling may impact borehole temperature reconstructions because the past SAT long-term variations could not be fully contained in the subsurface temperature structure.

Apart from the physical processes that interrupt the long-term SAT–GST coupling, the borehole reconstructions at large spatial scales might be further affected by a set of methodological aspects. These issues are represented by the sparse spatial distribution of borehole logs and especially by the uneven timing of the global borehole network's records. Accordingly, the CESM-LME was subsequently used to develop a pseudo-proxy experiment. The results indicate that borehole reconstructions are indeed able to retrieve the past GST long-term variations under an idealized scenario in which an appropriate spatiotemporal coverage of borehole logs is available. Nevertheless, under real-world conditions, the aggregated effect of the physical SAT–GST decoupling and the methodological constraints can lead to an underestimation of SAT as large as 50% of the total temperature increase within the 20th century on a global scale in the model

world. The level of impact is, however, highly dependent on the realization of internal variability. Intriguingly, the combined effect of the methodological and physical bias to borehole reconstructions yield an underestimation of the warming over the industrial period. This is relevant for climate reconstructions of LM temperatures and understanding uncertainties in pre-industrial to present temperature changes.

Publications related to this Thesis

The majority of the findings of this Thesis are included in the following publications.

- Melo-Aguilar, C., González-Rouco, J. F., García-Bustamante, E., Navarro-Montesinos, J., Steinert, N., 2018: Influence of radiative forcing factors on ground–air temperature coupling during the last millennium: implications for borehole climatology. *Climate of the Past*, **14**(11), 1583–1606. DOI: 10.5194/cp-14-1583-2018

- Melo-Aguilar, C., González-Rouco, J. F., García-Bustamante, E. Steinert, N., Jungclaus, J. H., Navarro, J., Roldán-Gómez, P. J, 2020: Methodological and physical biases in global to subcontinental borehole temperature reconstructions: an assessment from a pseudo-proxy perspective. *Climate of the Past*, **16**(2), 183-205. 10.5194/cp-16-453-2020

- Melo-Aguilar, C., et al.: Shallow soil thermal regime and surface air-ground temperature coupling over peninsular Spain: an assessment using a daily soil temperature dataset for Spain. *International Journal of Climatology.* (In preparation)

Additionally, during the development of this thesis the author has also collaborated in other published research articles:

- Roldán-Gómez, P. J., González-Rouco, F., Melo-Aguilar, C., Smerdon, J.: Dynamical and hydrological changes in climate simulations of the last millennium., Climate of the past Clim. Past, 16, 1285–1307, 2020, DOI: 10.5194/cp-16-1285-2020

- Fredrik Charpentier Ljungqvist, Andrea Seim, Paul J. Krusic, Fidel, González-Rouco, Johannes P Werner, Edward R Cook, Eduardo Zorita, Jurg Luter-

bacher, Elena Xoplaki, Georgia Destouni, Elena García-Bustamante, Camilo Melo-Aguilar, Kristina Seftigen, Jianglin Wang, Mary H Gagen, Jan Esper, Olga Solomina, Dominic Fleitmann and Ulf Bungen.,: European warm-season temperature and hydroclimate since 850 CE., Environmental Research Letters, 14, 8, doi: 10.1088/1748- 9326/ab2c7e, 2019.

Interacciones tierra–aire y el transporte de calor en el subsuelo: el papel de los forzamientos externos e implicaciones para la reconstrucción de temperaturas del último milenio

Los intercambios de energía en la interface suelo-atmósfera juegan un papel fundamental en numerosos procesos meteorológicos y climáticos tanto en la atmósfera como bajo la superficie. Una parte de la energía que impacta la superficie es reflejada o disipada mediante flujos de calor sensible y latente mientras que el exceso de energía incidente es balanceado a través del flujo de calor hacia el suelo. Este balance de energía en la superficie establece un vínculo entre la atmósfera y la superficie terrestre que generalmente se traduce en un fuerte acoplamiento entre la temperatura superficial del aire (SAT por sus siglas en inglés) y la temperatura de la superficie del suelo (GST por sus siglas en inglés). A su vez, la temperatura de la superficie se establece como la variable fundamental, ya que afecta a la cantidad de energía que fluye hacia el suelo mediante conducción térmica, por lo que influencia la distribución de temperaturas en el suelo. El acoplamiento entre SAT y GST, sumado al transporte de energía conductivo de la temperatura superficial hacia el suelo, representan las hipótesis centrales para reconstruir la temperatura superficial del pasado usando perfiles de temperatura del subsuelo (BTP por sus siglas en ingleś) actuales. Al igual que todos los métodos de reconstrucción de la temperatura, esta técnica también esta sujeta a fuentes de incertidumbre, que principalmente se dan por la influencia de algunos procesos físicos en la superficie que interrumpen el acoplamiento entre SAT y GST a diferentes escalas temporales. Entender los mecanismos físicos detrás de este tipo de procesos es crucial

para ganar confianza en las estimaciones de las variaciones en la temperatura del pasado.

Objetivos del estudio

El objetivo principal de esta Tesis es mejorar el conocimiento de las interacciones entre la superficie del suelo y la atmósfera que influencian el acoplamiento entre SAT y GST a diferentes escalas temporales y espaciales. Además, este trabajo también busca evaluar el transporte de energía en el subsuelo. Los supuestos de que hay un fuerte y estable acoplamiento entre el suelo y aire y que el régimen térmico en el subsuelo es conductivo, son las hipótesis centrales que permiten recuperar los cambios de la temperatura en superficie del pasado usando BTPs. Para tratar estos temas, se han usado tanto datos de observaciones como de modelos de clima. Se desarrollo una base de datos de registros de temperatura del suelo para España, la cual, por si misma, representa una contribución dada la escasez de este tipo de productos. Esta permite evaluar el transporte de energía en el subsuelo, así como el acoplamiento entre SAT y GST en escalas interanuales y decadales, considerando diferentes tipos de suelo y condiciones climatológicas. Además, se utiliza un conjunto reciente de simulaciones del último milenio (LM por sus siglas en inglés), producidas con el modelo del sistema tierra "Community Earth System Model" (CESM-LME), para explorar los mecanismos físicos que controlan los flujos de energía en la superficie en escalas temporales multicentenales. Así mismo, el uso de este conjunto de simulaciones permite entender los efectos individuales de los forzamientos externos del LM sobre las interacciones suelo-atmósfera y sus implicaciones para el acoplamiento SAT–GST. Esto no ha sido abordado anteriormente en simulaciones milenarias en las que el efecto de cada forzamiento externo es considerado.

Finalmente, este trabajo tiene como objetivo evaluar el desempeño de la técnica de reconstrucción de temperaturas a partir de BTPs, para recuperar las variaciones de largo plazo a grandes escalas espaciales durante el LM bajo un enfoque pseudo-proxy. Esto se centra en cuantificar los efectos debidos a una serie de limitaciones metodológicas, que incluyen la disponibilidad espacial y temporal de BTPs a nivel mundial, las cuales pueden afectar a la recuperación de GST del pasado. Adicionalmente, se explora la influencia de los procesos físicos en el desacoplamiento a largo plazo entre el SAT y GST sobre las estimaciones de las variaciones de SAT del pasado a través de BTPs. Esto no ha sido estudiado sistemáticamente antes.

Resultados más relevantes y conclusiones

Se ha compilado una base de datos de 39 estaciones que monitorean la temperatura del suelo a los niveles de 0.05, 0.1, 0.2, 0.5 y 1 m de profundidad. A esta base de datos se ha aplicado un control de calidad y representa una contribución útil para la comunidad científica interesada en las temperaturas del suelo y las interacciones suelo-aire. El análisis de esta base de datos ha mostrado que el transporte de energía en el primer metro del subsuelo es predominantemente conductivo, a pesar de otros mecanismos de transferencia de energía que también influyen en el reǵimen térmico del subsuelo. Así mismo, se ha evidenciado que la conexión entre SAT y GST es generalmente fuerte en España. Sin embargo, debido al efecto de algunos procesos estacionales, como el calentamiento radiaitvo de la superficie del suelo en los meses de verano, sumado a los efectos de la vegetación y la humedad del suelo, generan una gran diferencia entre SAT y GST que altera el supuesto del acoplamiento. Además, se ha visto que en escalas decadales se puede desarrollar un desacoplamiento entre SAT y GST sobre España central, debido a un mayor calentamiento de la superficie del suelo con respecto al aire durante las últimas décadas. La principal razón para esto puede ser la disminución de la precipitación y de la humedad superficial, que limita la cantidad de agua disponible en el suelo especialmente en los meses de verano. Al haber menos agua en el suelo para evaporar, la alta incidencia de radicación solar calienta la superficie del suelo, derivando en una respuesta general de mayor calentamiento de la superficie del suelo respecto al aire.

Un importante aspecto de esta Tesis es evaluar si el acoplamiento entre SAT y GST puede ser también interrumpido en escalas temporales multi-centenales. El conjunto de simulaciones CESM-LME muestra que globalmente y a escalas temporales multi-decadales y de menor frecuencia, las variaciones de SAT y GST están fuertemente acopladas durante el LM. Sin embargo, en escalas regionales existen varias situaciones que pueden determinar un desacoplamiento a largo plazo entre SAT y GST, debido a la influencia de los forzamientos externos. Esto incluye modificaciones en los flujos de energía en la interface suelo-atmósfera debido a cambios en la cobertura y uso del suelo. Así mismo, las interacciones entre el calentamiento por los gases de efecto invernadero y otros forzamientos externos, generan retroalimentaciones a largo plazo en la cobertura de nieve que pueden derivar en un desacoplamiento entre SAT y GST durante el periodo industrial. Estos procesos a escalas locales y regionales pueden impactar las reconstrucciones mediante el uso de BTPs ya que la señal de SAT del pasado no estaría completamente contenida en la estructura de temperaturas del subsuelo.

Aparte de estos procesos físicos, que interrumpen el acoplamiento entre SAT y GST, las reconstrucciones basadas en BTPs a grandes escalas espaciales pueden

sufrir de otro tipo de sesgos; debido a aspectos metodológicos que están representados por la dispersa distribución espacial de BTPs, especialmente por la variabilidad en la fecha de registro existente en la red global de BTPs disponibles. De acuerdo a esto, se ha usado el conjunto de simulaciones del CESM-LME para el desarrollo de un experimento pseudo-proxy. Los resultados muestran que, en condiciones ideales, en las que se tuviera una cobertura apropiada de BTPs registrados en una fecha homogénea, esta técnica de reconstrucción es capaz de reproducir las variaciones de largo plazo de la GST del pasado. Sin embargo, en las condiciones del mundo real el efecto del desacoplamiento entre SAT y GST, sumado a las limitaciones metodológicas, puede generar una infraestimación en el aumento de la temperatura del aire durante el siglo XX a escala global, de hasta un 50% en la realidad sustituta del modelo. Curiosamente, el efecto combinado de los aspectos metodológicos y físicos conllevarían a una infraestimación del calentamiento industrial. Esto es relevante para reconstrucciones climáticas del LM y para entender incertidumbres en los cambios de temperatura desde el periodo preindustrial hasta el presente.

Publicaciones relacionadas con esta Tesis

La mayoría de los resultados de esta Tesis se incluyen en las siguientes publicaciones.

- Melo-Aguilar, C., González-Rouco, J. F., García-Bustamante, E., Navarro-Montesinos, J., Steinert, N., 2018: Influence of radiative forcing factors on ground–air temperature coupling during the last millennium: implications for borehole climatology. *Climate of the Past*, **14**(11), 1583–1606. DOI: 10.5194/cp-14-1583-2018

- Melo-Aguilar, C., González-Rouco, J. F., García-Bustamante, E. Steinert, N., Jungclaus, J. H., Navarro, J., Roldán-Gómez, P. J, 2020: Methodological and physical biases in global to subcontinental borehole temperature reconstructions: an assessment from a pseudo-proxy perspective. *Climate of the Past*, **16**(2), 183-205. 10.5194/cp-16-453-2020

- Melo-Aguilar, C., et al.,: Shallow soil thermal regime and surface air-ground temperature coupling over peninsular Spain: an assessment using a daily soil temperature dataset for Spain. *International Journal of Climatology.* (In preparation)

Adicionalmente, durante el desarrollo de esta tesis el autor también participo en otras publicaciones científicas:

- Roldán-Gómez, P. J., González-Rouco, F., Melo-Aguilar, C., Smerdon, J.: Dynamical and hydrological changes in climate simulations of the last millennium., Climate of the past Clim. Past, 16, 1285–1307, 2020, DOI: 10.5194/cp-16-1285-2020

- Fredrik Charpentier Ljungqvist, Andrea Seim, Paul J. Krusic, Fidel, González-Rouco, Johannes P Werner, Edward R Cook, Eduardo Zorita, Jurg Luterbacher, Elena Xoplaki, Georgia Destouni, Elena García-Bustamante, Camilo Melo-Aguilar, Kristina Seftigen, Jianglin Wang, Mary H Gagen, Jan Esper, Olga Solomina, Dominic Fleitmann and Ulf Bungen.,: European warm-season temperature and hydroclimate since 850 CE., Environmental Research Letters, 14, 8, doi: 10.1088/1748- 9326/ab2c7e, 2019.

List of acronyms

AEMET	Agencia Estatal de Meteorología/Spanish Meteorological Agency
ALL-F	All Forcings Simulation
BBCP	Bottom Boundary Condition Placement
Bmask	Borehole Masked Scenario
BTP	Borehole Temperature Profile
CE	Common Era
CESM	Community Earth System Model
CESM–LME	Community Earth System Mode–Last Millennium Ensemble
CLM4	Community Land Model version 4
CMIP5	Climate Model Intercomparison Project Version 5
DJF	December–January–February
E	East
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA5-L	ERA5-Land Reanalysis
ESM	Earth System Model
GCM	General Circulation Model
GHGs	Green House Gases
GMT	Greenwich Mean Time
GST	Ground Surface Temperature
IBS	Ideal Borehole Scenario
JJA	June–July–August
LHFLX	Latent Heat Fluxes
LIA	Little Ice Age
Ln-A	Natural Logarithm of the Amplitude
LM	Last Millennium
LULC	Land Use Land Cover

MAM	March–April–May
MCA	Medieval Climate Anomaly
N	North
NH	Northern Hemisphere
OZ/AER	Anthropogenic Ozone/Aerosols
PCs	Principal Components
PMIP3	Paleoclimate Model Intercomparison Project version 3
PPE	Pseudoproxy experiment
P-S	Phase Shift
QC	Quality Control
RMSE	Root Mean Square Error
RSW	Reflected Shortwave Radiation
S	South
SAT	Surface Air Temperature/2m air temperature
SIC	Soil Ice Content
Singe-F	Single Forcing simulation
SH	Southern Hemisphere
SHFLX	Surface Sensible Heat Flux
SON	September–October–November
SoTOS	Soil Temperature Observations for Spain
SSW	Surface Shortwave Radiation
ST	Soil Temperature
SVD	Singular Value Decomposition

1

Introduction

1.1 Land–atmosphere coupling

The land–atmosphere interactions via the energy and water exchanges at the ground surface play a fundamental role influencing several weather and climatic processes at different spatial and temporal scales. A part of the radiation that reaches the surface is reflected away, i.e albedo, or dissipated by sensible and latent heat fluxes back to the atmosphere providing the source of energy and moisture to drive the hydrological cycle. These energy fluxes are modulated by the characteristics of the land surface such as the vegetation cover, soil moisture content, soil type and others (Geiger et al., 1995; Beltrami and Kellman, 2003). The changes in the incoming energy are balanced by long wave emission and by the ground heat flux, therefore influencing heat storage in the subsurface. Besides the recognized effects on climate, the surface energy balance establishes a link between the atmosphere and the ground, with the surface temperature as the key variable representing equilibrium.

The physical interactions at the Earth's surface result in a coupling between the state of the atmosphere and the land (Bonan, 2015) that generally translates into a strong connection between the surface air temperature (SAT) and the ground surface temperature (GST). The surface temperature affects the amount of heat flowing into the soil, thus controlling the distribution of temperature below the surface. Accordingly, two interesting aspects regarding the thermodynamics of the subsurface emerge from the interactions between the land and the atmosphere. First, the temperature at the surface is coupled to the air temperature

due to the energy balance at the ground-atmosphere interface, and second, a climatic signal of the temperature variations at the surface propagates downwards into the subsurface by thermal conduction. If these two hypotheses are supported, this will determine the thermal state below the surface and its implications for past climate reconstructions.

1.2 The subsurface thermal regime

The study of the energy transport below the surface has been an important field of research in the geothermal science because understanding the subsurface thermal regime is crucial in the analysis of the several geophysical processes, as for instance, the terrestrial heat flow determination or heat storage in the subsurface (Bodri and Cermak, 2007b). The distribution of temperature in soils is also important for several climate-related and ecological studies. This is the case, for example, in assessments of carbon dioxide (CO_2) release from soils to the atmosphere due to soil respiration processes since soil temperature is one of the main factors controlling both plant root respiration and microbial respiration (Carey et al., 2016). In particular, the response of soil respiration to increase in temperatures under global warming conditions has raised considerable attention because the transfer of carbon from soils to the atmosphere would positively feedback climate change (Melillo et al., 2002; Davidson and Janssens, 2006; Bond-Lamberty and Thomson, 2010). Likewise, the soil temperature is a key variable for agricultural oriented studies that can influence root and plant growths (Kaspar and Bland, 1992; McMichael and Burke, 1998) as well as crops yieldings of some grains (Cooper and Law, 1978; Zhang et al., 2019).

Owing to the relevance of soil temperature on controlling a variety biophysical and biochemical processes in soils, it is important to understand its temporal and spatial variability as well as the response to atmospheric warming (Jungqvist et al., 2014). *In situ* soil temperature measurements at different depths in meteorological stations represent the most adequate source of information to evaluate the distribution of temperature in soils and to address its influence on soil biological and chemical processes as well as on climate feedbacks. Although soil temperature observations are scarce both in space and time since this variable has received less attention than other meteorological variables (e.g., air temperature and precipitation; Qian et al., 2011), during the recent decades there have been some advances in the development of soil temperature databases with local or regional coverage around the globe. For instance, Hu and Feng (2003) developed a soil temperature database for the U.S. including five soil layers down to

a depth of 1 m. Likewise, the U.S. Climate Reference Network (USCRN) has incorporated soil temperature measurements at various depths since 2011 (Bell et al., 2013) and the Soil Climate Analysis Network (SCAN) includes soil temperature records since the late 20th century also for the U.S. (Schaefer and Paetzold, 2000). Some regional soil temperature observational datasets have also been developed in other parts of the world such as Canada (Qian et al., 2011), China (Zhang et al., 2016) and the northern high latitudes in Russia (Zhang, 2000). In the recent years the use of earth observation systems to estimate temperature below the surface is becoming an opportunity to expand the soil temperature observations network (Huang et al., 2020).

Despite the current developments in soil temperature datasets around the globe, there are still limitations in the spatial and temporal availability of soil temperature data (Hao et al., 2014; Huang et al., 2020) as systematic observations are only available since the mid-20th century in the best cases, mostly distributed in the northern mid and high latitudes. Therefore, expanding the spatial and temporal coverage of soil temperature records is of great interest to promote analyses regarding the soil thermodynamics and the response to atmospheric warming over areas with different soil properties, surface characteristics and climatological conditions. Hence, the development of additional soil temperature databases in other parts of the world, either from *in situ* measurements or remote sensing information, is highly desirable. Considering this need, soil temperature observations over peninsular Spain and the Balearic Islands, down to a depth of 1m, have been gathered and quality controlled in this Thesis to create a soil temperature observational dataset for Spain (SoTOS; cf. Chapter 2). This will allow for insights on the regional/local land-air interactions and thermal state below the surface.

Indeed, the use of these soil temperature datasets has allowed for exploring different aspects regarding the soil thermodynamics and its effects on soil biological and chemical processes. For instance, Hu and Feng (2003) analyzed the soil temperature evolution over the U.S. during the last half of the 20th century. They found that soil temperatures within the upper meter below the surface have steadily increased during this period in most of the analyzed sites over the U.S. with an average rate of $\sim$0.31 °C per decade. Similarly, they argue that such a warming trend in soil temperatures may have both positive and negative impacts on agriculture as it would favor sowing earlier in the year, thus enhancing the growth of some crops that require longer seasons to mature. However, it could also benefit the survival of insects in winter and increase populations during the planting season that may harm the development of crops. Analyses in Canada (Qian et al., 2011) and China (Zhang et al., 2016) have also shown rising temperatures in soils during the late 20th century and the beginning of the 21st century,

even at relatively deep soil layers (i.e., 1.5 to 3.2 m depth) with the largest increases at different seasons depending on the geographical distribution and the climatology of the locations included in both cases. Zhang et al. (2016) reported that warming in soils is positively correlated to increased soil respiration and also to larger growing seasons for agricultural production in China.

The assessment of the downward propagation of the surface temperature changes has also received considerable attention due to its relevance for some branches of climate research. Particularly, for paleoclimate analyses as subsurface temperatures store a signal of the long-term past surface climate variations that are imprinted in it as their signal penetrates (Shen and Beck, 1991). Due to the nature of heat conduction, the surface temperature changes slowly propagate into the subsurface as a function of the timescale such that the soil acts as a low-pass filter. As a consequence, the higher frequency components are rapidly filtered out and only detectable within the few upper meters below the surface while the low–frequency changes penetrate deeper into the subsurface (Mareschal and Beltrami, 1992). The latter can be clearly seen in the response of soil temperatures to the propagation of surface temperature variations with time as represented by continuous measurements in meteorological observatories at different soil layers. Figure 1.1 shows the temperature evolution at a site located at the Sierra de Guadarrama, Spain, belonging to the Guadarrama Monitoring Network (GuMNet*; Vegas-Cañas et al., 2020). It is evident that soil temperatures do not simply vary synchronously with the surface signal. Rather, as the signal penetrates, the high-frequency changes (e.g., daily or synoptic variations) are progressively filtered out and almost completely absent at a depth of 1 m. In turn, the annual cycle propagates more in depth and can be detected within the upper 5 m, but it is very attenuated below 10 m depth. Indeed, for this type of temperature variations that can be approximated to a periodic function of time (e.g., daily or seasonal cycles), the conductive heat transport implies an exponential amplitude decay and a linear phase shift with depth (Turcotte and Schubert, 2014). Such characteristics are also evident in Figure 1.1. The rate of amplitude attenuation and phase shift of the soil temperatures relative to the surface signal depends on the thermal properties of the subsurface that are controlled by the mineral and organic composition of the soil and the water content that ultimately determine the thermal diffusivity (Cosenza et al., 2003). In the shallow subsurface, where variations in the water table with time exist, the thermal diffusivity can vary significantly with time due to its dependence on soil moisture (Ochsner et al., 2001). On the contrary, deeper in the bedrock, the thermal diffusivity can be assumed to be constant, as variations with time in the subsurface properties are negligible (Chisholm and Chapman, 1992).

* www.ucm.es/gumnet/

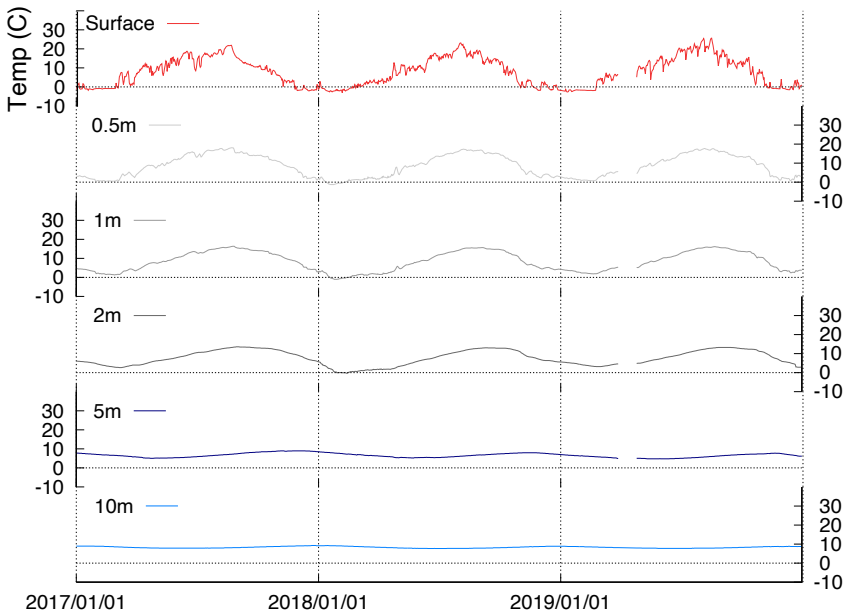

Figure 1.1: Daily surface and soil temperature at 0.5, 1, 2, 5 and 10 m depths measured at the Cotos site in the Sierra de Guadarrama, Spain, in a 20 m deep borehole over the 2017–2020 period, belonging to the GuMNet facility (Vegas-Cañas et al., 2020).

The characteristics of conductive downward heat transport of the surface temperature changes have been explored by several works with the use of soil temperature observations. Smerdon et al. (2003) employed soil temperature data up to 11 m depth at Fargo, North Dakota, to illustrate that thermal conduction effectively represents the distribution of temperature with depth as the analyzed soil layer exhibited the exponential amplitude attenuation and linear phase shift with the depth of the annual cycle. Further, they estimated the mean thermal diffusivity for the soil profile at Fargo as a metric of the subsurface conductive regime based on the linear dependency of the amplitude attenuation and phase shift with depth. This analysis was extended to other few locations over the U.S. and the Czech Republic by Smerdon et al. (2004), showing that the basic characteristics of the conductive propagation of the surface temperature into the soil also represent the temperature distribution with depth over a variety of soil types. Some other studies have also used subsurface temperatures over a few locations

to analyze the conductive heat transport in soils, yielding similar evidence of the prevalence of thermal conduction on controlling the distribution of temperature in the subsurface (e.g., Putnam and Chapman, 1996; Schmidt et al., 2001; Baker and Baker, 2002; Cermak et al., 2017).

Even though conductive heat transport of surface temperature changes is the main mechanism controlling the depth and temporal variability in soil temperatures, subsurface temperature profiles recorded in deep boreholes have shown that other heat transport mechanisms can also influence the temperature distribution below the surface. This is the case of vertical and horizontal subsurface fluid movements (hydraulic advection), effects of changes in terrain conditions (erosion or burial) or convection (Lewis and Wang, 1992; Chisholm and Chapman, 1992). Over some areas, these mechanisms can be the most important factor responsible for the temperature variations with depth, especially vertical and horizontal groundwater movements where the permeability structure of the bedrock is high (Kukkonen and Clauser, 1994). Nevertheless, it is widely recognized that in most situations, the conductive heat transport of the surface temperature variations dominates the distribution of temperature within the subsurface (Kohl, 1998; Bodri and Cermak, 2005; Ferguson et al., 2006). In this Thesis, the soil temperature data from SoTOS is employed to expand previous analysis regarding the temperature propagation into the shallow subsurface. Chapter 3 evaluates the characteristics of the conductive heat transport over a variety of sites over Spain with different soil physical properties and climatic conditions.

1.3 Surface processes influencing the land-air temperature coupling

The use of soil and air temperature observations has also allowed the development of assessments regarding the interactions between the atmosphere and the ground surface. This is important because the bond between SAT and GST may be disrupted by changes in the land surface characteristics that alter the energy balance. Therefore, understanding the mechanisms that modulate the energy fluxes at the surface is relevant to evaluate situations in which the connection between SAT and GST is disrupted. These types of processes may occur at a variety of timescales from daily and seasonal cycles to decadal, centennial and longer-term variations as a response to external forcings (Melo-Aguilar et al., 2018). In such cases, the temperature signal that propagates into the subsurface may be uncorrelated to the temperature variations in the overlying atmosphere.

One of the most important factors that play some role in modifying the net energy exchanges between the ground and the atmosphere is represented by the seasonal snow cover. As snow cover is a strong insulator, the exchange of heat between the atmosphere and the ground is limited when significant amounts of winter snow cover are present, thus leading to large offsets between SAT and GST (Pollack and Huang, 2000; Beltrami and Kellman, 2003; Stieglitz et al., 2003; Bartlett et al., 2004). Similarly, seasonal changes in soil moisture and evapotranspiration due to land management practices such as irrigation result in different responses of SAT and GST with a general pattern of lower soil temperature relative to the air above due to evaporative cooling (Wang et al., 2000). This is because the incoming energy that reaches the Earth's surface is mostly employed to evaporate the available source of water in the soil, thus leading to a net heat loss at the ground surface (Melo-Aguilar et al., 2018). On longer timescales, processes associated with land use land cover (LULC) changes also have the potential to impact the coupling between SAT and GST. This is mainly due to the influence of biogeophysical mechanisms like changes in the physical properties such as albedo, roughness and evapotranspiration that modify the surface energy balance (Pongratz et al., 2010; Anderson et al., 2011). For instance, deforestation over mid and high latitudes can lead to atmospheric cooling due to an increase in albedo as open fields have a higher reflectivity than forested areas (Jackson et al., 2008). The latter also has an indirect effect on soil temperatures as deforestation at these latitudes can enhance winter snow cover (Claussen et al., 2001; Brovkin et al., 2006), thus leading to the warming of the ground surface due to the insulating effect of snow cover. In lower latitudes, deforestation also leads to different responses in SAT and GST (Winckler et al., 2019). Figure 1.2 illustrates the variations in the energy fluxes at the surface due to different LULC transitions from forested areas to open fields over different geographical locations. The effects are shown for winter and boreal summer scenarios to highlight different seasonal processes and associated feedbacks, such as enhanced winter snow cover. In general, deforestation reduces the transport of heat from the surface to the atmosphere by different mechanisms depending on the geographical location, thus impacting the connection between SAT and GST even on centennial timescales (MacDougall and Beltrami, 2017). Accordingly, processes of diverse nature acting at different timescales can impact the connection between the land and the atmosphere, thus influencing the signal that ultimately propagates into the subsurface.

Some of these issues have been addressed in the literature by taking advantage of the available surface air and soil temperature observational records that permit the analysis of the SAT–GST interactions over a variety of land surface and soil characteristics. Particularly, the role of snow cover has been broadly investigated ought to its large potential to interrupt the connection between the

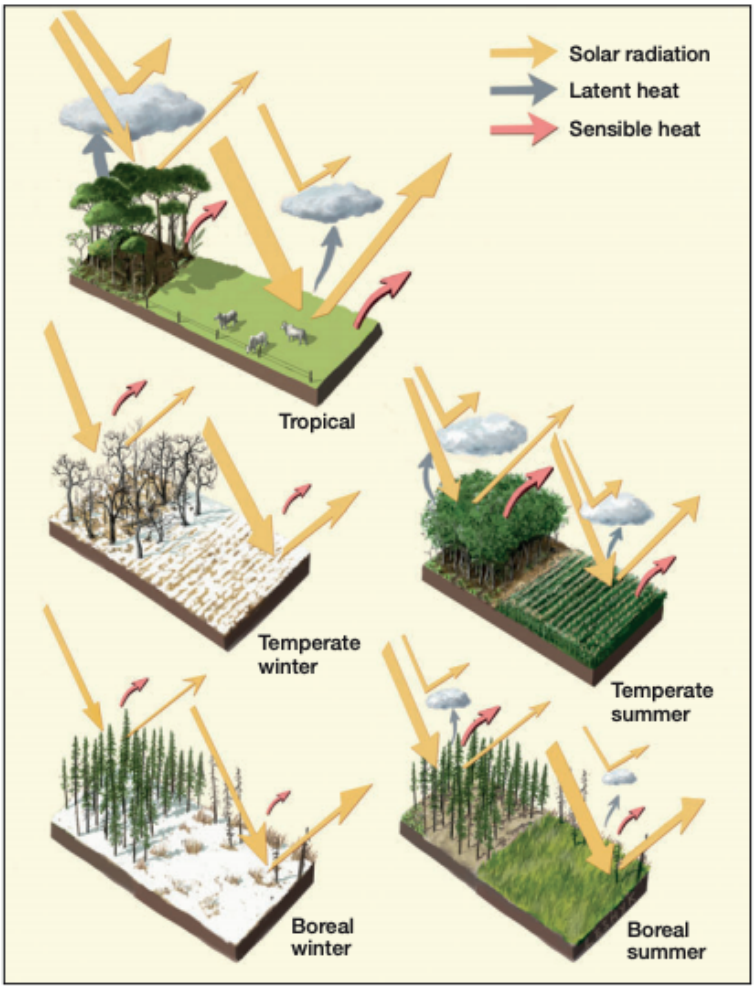

Figure 1.2: Changes in the surface energy fluxes from forest to non-forest transitions in tropical, temperate and boreal ecosystems. The effects are shown for winter and boreal summer scenarios (from Anderson et al., 2011).

atmosphere and the ground surface. Smerdon et al. (2003) showed that large winter SAT–GST offset exists over areas with a large accumulation of snow during winter and spring seasons. Likewise, It has also been described that changes in the onset, duration and amount of snow cover in the coldest months influence

the propagation of the air temperature signal into the ground and can lead to variations in the SAT–GST relationship at decadal timescales (Bartlett et al., 2005). Other processes acting at short timescales can also have an effect on interrupting the connection between the air and ground temperature. For instance, the radiative heating of the ground surface during the summer months results in an overall response of larger GST warming relative to the SAT. In locations with low rates of soil moisture content during the summer months, when solar radiation reaching the surface is largest, this effect can significantly impact the SAT–GST coupling at seasonal timescales (Putnam and Chapman, 1996). Similarly, Smerdon et al. (2004) further demonstrate that other seasonal processes, as for instance, summer precipitation and evapotranspiration rates, may also interfere in the SAT–GST coupling at short timescales due to latent heat exchanges. The effects of short-term processes on the SAT–GST coupling have been further explored by several studies using observational data (e.g., Beltrami and Kellman, 2003; Smerdon et al., 2006; Cermak and Bodri, 2018). Overall, these works have found that the energy exchanges at the air-ground interface due to several processes at seasonal timescales such as latent heat uptake and release due to some hydrological processes, winter snow cover or radiative heating, among others, can lead to short-term SAT–GST decoupling. Nevertheless, it has also been stated that such short-term processes would not affect the SAT-GST association on longer timescales if they are stationary as its influence would be a constant offset with time (Bartlett et al., 2005).

This Thesis addresses the relationship between SAT and GST over some locations in Spain with the use of air and soil temperature information from SoTOS. Chapter 3 presents an assessment of the SAT–GST coupling on interannual to multi-decadal timescales. This analysis provides additional insights into the land-air interactions that modulate the energy fluxes at the ground surface, and thus, the SAT and GST relationship at theses timescales.

1.3.1 Model simulations to assess the long-term land–air temperature coupling

The assessment of the complex SAT–GST relationship over longer timescales using observational data is limited by the relatively low availability of soil temperature observations for periods longer than a couple of decades, as discussed earlier. An alternative is using model simulations as a complementary source of information about land–air interactions.

The discussion of the connection between SAT and GST over long timescales has benefited considerably from introducing contributions from the modeling world that allowed for addressing the SAT–GST relationship at longer timescales than a couple of decades. In addition, the spatial coverage is not limited in this type of analysis. Mann and Schmidt (2003) used both air and soil temperatures simulated with the General Circulation Model (GCM) of the Goddard Institute for Space Studies (GISS modelE) that span from 1951 to 1998 of the Common Era (CE). They claimed that the seasonal influence of snow cover could have a large influence on decoupling SAT and GST since changing snow cover conditions with time may result in uncorrelated variations in GST with respect to SAT. Some other works have shown that SAT and GST coupling is strong at longer than seasonal timescales; therefore, snow biases would influence only the high-frequency oscillations. For instance, González-Rouco et al. (2003) investigated the relationship between simulated SAT and GST from interannual to centennial timescales in a forced climate simulation of the last millennium (LM) with the ECHO-G GCM (Legutke and Voss, 1999). They found that in spite of the seasonal and longer-term variability of snow cover, the coupling was stable at longer timescales and especially at large spatial scales. Thus, the GST should be a good proxy for long-term SAT variations on a global scale. Nevertheless, the climate models employed in these earlier works do not incorporate the variations in vegetation cover (Jansen et al., 2007), which is recognized as an important factor that can alter the energy exchanges at the land-atmosphere interface and thus impact the SAT–GST relationship.

Recently, a couple of studies have further investigated the SAT–GST relationship in model simulations with state-of-the-art Earth System Models (ESMs) within the frame of the CMIP5 (Coupled Model Intercomparison Project phase 5; Taylor et al., 2012) over historical and future climate scenarios for North America (García-García et al., 2019) and at global scales (Soong et al., 2020). Both of these works have shown that the SAT–GST relationship may experience variations in the 21st century as simulated by the current generation of ESMs since the rate of warming is higher for SAT compared to GST in North America and also for the global case. They reported that this effect is larger at the high northern latitudes and is mainly influenced by changes in snow cover during the 21st century. Even though these works have reported that SAT–GST decoupling may exist as a response to global warming, the specific drivers (i.e., external forcings) and the related feedbacks at either global/hemispherical or continental/regional scales were not addressed. This is because the model simulations employed in such analyses are developed in the so-called "all-forcing" framework, in which all of the natural and anthropogenic external forcings are included. Understanding the contribution of individual forcing factors is desirable in order to character-

ize the mechanisms with potential for modifying the land-atmosphere connection and thus perturbing the SAT-GST coupling.

Other modeling assessments have focused on specific mechanisms with the potential for changing the surface characteristics and soil properties and thus interrupt the SAT–GST coupling. Particularly, the LULC transitions from forested areas to open field or vice versa have received some attention as they have the largest potential to interrupt the SAT–GST coupling due to the modification of the land–surface properties that alter the energy fluxes in the ground-atmosphere interface (Anderson et al., 2011). The effects of deforestation were addressed by MacDougall and Beltrami (2017) with the use of a climate model of intermediate complexity in a controlled experiment in which deforestation is imposed progressively at a specific rate over each time step of the simulation. They found that deforestation tends to warm the ground surface mainly by reducing the transport of heat away from the surface. Such type of vegetation changes would drive to long-term surface temperature anomalies being uncorrelated to changes in the atmospheric conditions, and thus, deforestation should be considered as a source of bias to the strong connection between the state of the atmosphere and the land surface. Additional assessments of ESM simulations considering the individual effect of LULC changes under a realistic representation of past LULC as well as plausible realizations in the future (Pongratz et al., 2008; Hurtt et al., 2016) are important in order to improve our understanding of the SAT and GST responses to changes in the land surface characteristics including transitions other than deforestation. Further evaluations from the modeling world of the influence of LULC changes on corrupting the long-term relationship between SAT and GST were developed by Melo-Aguilar et al. (2018).

This Thesis includes the assessment of Melo-Aguilar et al. (2018) in which LM simulations from the Community Earth System Model–LM Ensemble (CESM–LME; Otto-Bliesner et al., 2016) were employed to further investigate the long–term relationship between SAT and GST. Particularly, the effects of individual forcing factors on the long-term SAT–GST coupling over the LM was addressed. This was done by using the single-forcing simulations (i.e., simulations forced by individual external forcings) included in the CESM–LME. These single-forcing experiments are suitable for identifying the specific role that each forcing might play and its "fingerprint" on producing long-term variations in the SAT–GST relationship. The details of the CESM-LME are described in Chapter 2 and the findings of Melo-Aguilar et al. (2018) are presented in Chapter 4.

1.4 Borehole-based temperature reconstructions

The conductive–dominated heat transport of the temperature changes at the surface into the subsurface, in addition to the interconnected state of the atmosphere and the ground surface, have allowed the development of a whole branch of paleoclimate research known as borehole climatology that focuses on the extraction of long–term past surface temperature signals from present-day borehole temperature profiles (BTPs). The history of this field can be traced back to the 1920s with the first work in this direction. Back then, the variations in the rate of increase of temperature with depth, estimated from temperature records in copper mines in the southern margin of Lake Superior, were associated with glaciation–deglaciation transitions occurring some 10,000 years ago (Lane, 1923). However, at the time, the main focus was to identify the origin of the difference in geothermal flow at different depths rather than to provide an accurate estimate of past surface temperature variations. The analyses have been refined with time and bloomed in the 1980s when several independent temperature reconstructions were developed at local scales over different locations around the globe (Bodri and Cermak, 2007b).

The borehole temperature reconstruction theory states that a signal of the past surface temperature variations is captured down below the surface as a result of conductive heat propagation. The nature of the phenomenon is as follows. The Earth's crust can be assumed to be in thermal equilibrium for the climatological timescales of interest (Bodri and Cermak, 2007b). Therefore, the heat flow from the Earth's interior results in a constant geothermal gradient characterized by a steady temperature increase with depth. Under this quasi-steady geothermal state, the temperature changes at the surface, that slowly propagate into the subsurface, are superimposed on the background geothermal gradient, thus curving the upper part of the temperature depth profile; positive/negative temperature depth anomalies correspond to surface warming/cooling (Beltrami and Bourlon, 2004). The downward propagation of the surface temperature changes is governed by the low thermal diffusivity of the rock, and the timescale of the temperature change such that the soil acts as a low–pass filter. As a consequence, the high-frequency surface variations cannot be detected below a few meters, whereas lower frequency oscillations reach deeper depths. Figure 1.3a shows a schematic representation of the depth penetration as a function of the timescale for a single surface temperature change of 1 °C occurring since 1, 10, 50, 100, 500 and 1000 years before present, assuming a thermal diffusivity of 1×10^{-6} m^2 s^{-1}, which is a typical value for rocks (Mareschal and Beltrami, 1992). Note that, gradually, the surface temperature change propagates to deeper depths as the duration of the temperature change is longer. For instance, multi-decadal or cen-

tennial trends and can be detected in the order of 100 to 300 m depth, whereas for millennial timescale changes, the subsurface perturbation reaches depths up to 500 to 650 m below the surface (Pollack and Huang, 2000). This frequency dependence of the conductive heat transport with depth determines one of the most distinctive features of borehole temperature reconstructions, namely, that only the low-frequency information (e.g., multi-decadal to centennial trends) can be retrieved from BTPs.

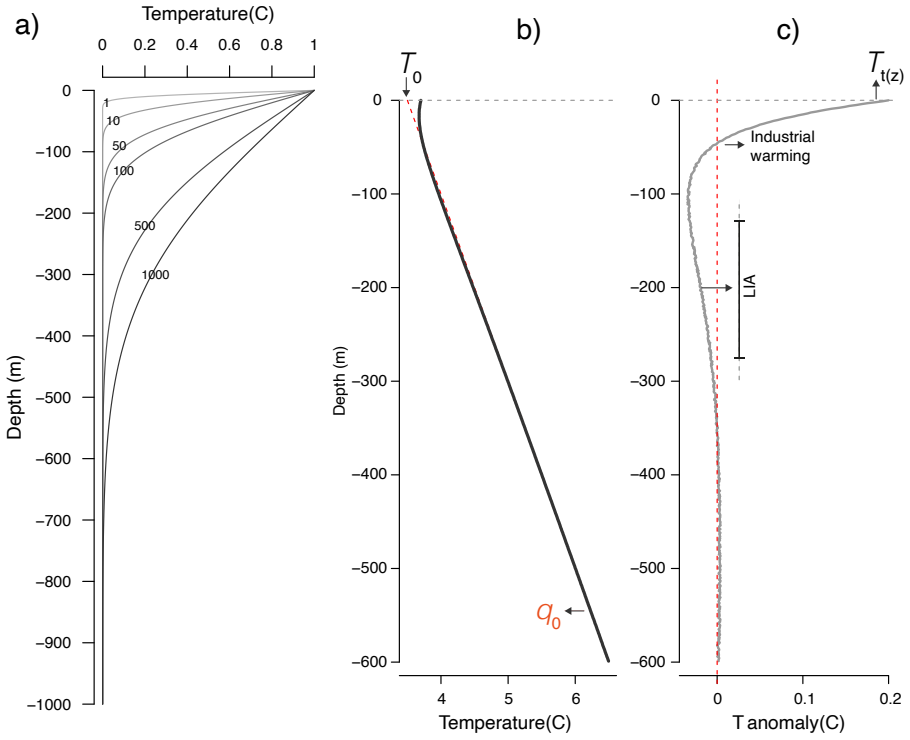

Figure 1.3: a) Subsurface temperature perturbation due to a single-step temperature change of 1 °C occurring since 1, 10, 50, 100, 500 and 1000 years before present (adapted from Mareschal and Beltrami, 1992). b) Schematic representation of a 600 m depth borehole temperature profile. c) Anomaly temperature profile after removing the background geothermal gradient.

In a BTP, the temperature distribution with depth, T (z), is thus composed by the superposition of the geothermal background field and the integrated signal of the surface temperature perturbations of various timescales:

$$T(z) = T_0 + q_0 R(z) + T_t(z), \tag{1.1}$$

where q_0 represents the surface heat flow density, $R(z)$ is the thermal depth, and T_0 is a reference ground temperature. Figure 1.3b illustrates a schematic representation of a BTP of 600 m (thick black line), including the different components of Equation 1.1. The quasi-steady geothermal state (dashed red line in Figure 1.3b) can be estimated from the bottom part of the BTPs, that is less affected by the recent past surface temperature variations, by linear fitting to the data and extrapolation to the surface. The resulting slope represents the geothermal gradient, while the interception with the surface stands as the equilibrium surface temperature. The background component is then subtracted from the BTP to extract the temperature anomalies associated with the transient perturbation component $T_t(z)$ that is shown in Figure 1.3c. The latter contains the information on the long-term past surface temperature variations, which is ultimately the target signal to be obtained from the BTPs. In the diagram shown in Figure 1.3c, a signal of a cold period sometime before the present is contained within the 100-300 m depth, as illustrated by the negative anomalies with respect to the background state. As an example, such a period can be associated with the period known as the little ice age (LIA), as the timescale of the depth range of the anomalies corresponds to a few hundred years. Similarly, a shift towards positive subsurface temperature anomalies at approximately 100 m depth represents the onset of industrial warming, with the largest temperature increases of the last decades contained in the upper 50 m of the reduced profile. To derive the long-term past surface temperature variations from the anomaly profile, different inversion techniques are usually employed, as for instance, least-squares inversion in functional space (Shen and Beck, 1991), singular value decomposition (Mareschal and Beltrami, 1992) or ramp changes (Pollack et al., 1996), under the assumption that the subsurface thermal state is a purely conductive process. The result is a reconstruction of the GST history that is limited to recover only the low-frequency variations due to the nature of heat conduction.

The use of subsurface temperature data to investigate the past evolution of surface temperatures can be dated back to the early 1970s with the first systematic studies to infer the past surface temperature variations from the underground temperature in individual boreholes (Cermak, 1971). Since then, the approach gradually gained some recognition and was more generally employed in the following decades to study the past evolution of the temperature at the surface

at different locations around the world (e.g., Lachenbruch and Marshall, 1986; Beltrami and Mareschal, 1991; Beck et al., 1992; Harris and Chapman, 1995; Huang et al., 1996). Thanks to the work of several independent research groups, a global dataset of borehole temperature measurements was gathered (Huang and Pollack, 1998), and borehole reconstructions at hemispheric to global scales started to be developed at the beginning of the 2000s. Several independent borehole reconstructions have indicated a surface temperature increase of ∼1 °C from 1500 to 2000 CE at a hemispheric to global scale (e.g. Huang et al., 2000; Harris and Chapman, 2001; Pollack and Smerdon, 2004; Beltrami and Bourlon, 2004; Masson-Delmotte et al., 2013).

1.5 Assessing borehole temperature reconstructions

Addressing the range of uncertainties in reconstructing past temperature changes is relevant not only for assessing our understanding of past temperature changes and the confidence we have in available estimates (Masson-Delmotte et al., 2013) but also for constraining the range of estimates of the system response to changes in the external forcing, i.e., the sensitivity of the climate system (Fernández-Donado et al., 2013). Figure 1.4 illustrates the set of temperature reconstructions of the CE for the northern hemisphere (NH) from a variety of proxy sources. An updated version of the borehole-based reconstruction of Pollack and Smerdon (2004) is represented by the thick red line that spans from 1500 to 2000 CE. The higher temperature increase of this reconstruction during the last five centuries compared to most other proxy reconstructions is remarkable. The estimates of past surface temperature evolution obtained from BTPs have been subjected to debate during the past few decades due to their reported multi-centennial trend of ca. 1 °C over the 1500–2000 CE period at hemispheric to global scale, a relatively large value in comparison with other proxy sources that estimate a temperature increase in the range of ∼0.6-0.8°C (Masson-Delmotte et al., 2013; Neukom et al., 2019). Therefore, BTPs are relevant sources of information to estimate the Earth System temperature response to external forcing (e.g., Fernández-Donado et al., 2013; Smerdon and Pollack, 2016).

The discussion regarding the reliability of the borehole reconstructions has been focused on whether the existence of physical factors with the potential to disturb the SAT–GST coupling at different temporal scales, as described in Section 1.3, may hinder the achievement of robust borehole-based estimations of past SAT long-term trends. Apart from the physical factors influencing the SAT–GST coupling assumption, assessments on the effects of some methodolog-

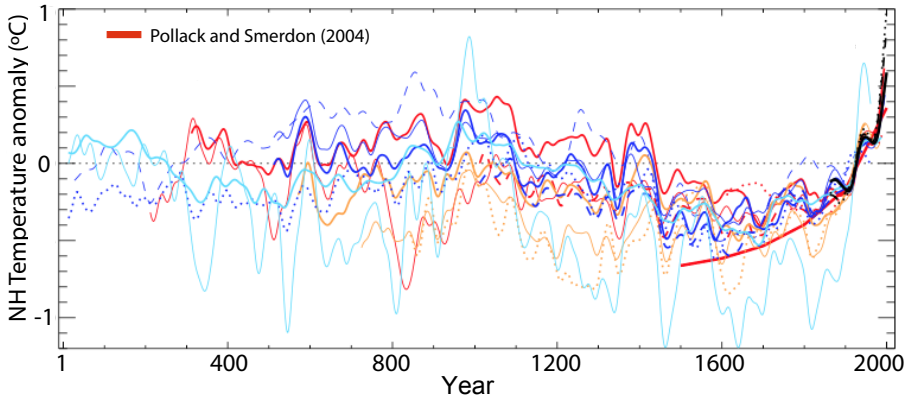

Figure 1.4: Northern hemisphere proxy-temperature reconstructions for the last 2000 years (colored lines) and instrumental temperature observational (black line), anomalies with respect to the 1881–1990 mean. Adapted from Masson-Delmotte et al. (2013).

ical constraints on the overall performance of the borehole estimates have also been carried out (e.g., Rutherford and Mann, 2004; Pollack and Smerdon, 2004). Methodological aspects refer to a variety of issues that include: site-specific processes contributing to noise in BTPs, interactions with orography and hydrology leading to horizontal (vertical) advection (convection) that render the conductive regime assumption invalid, sampling irregularities such as low density and/or uneven spatial distribution of boreholes at larger spatial scales, variability in logging dates and depth of profiles, errors in measurements, etc. (Bodri and Cermak, 2007a; González-Rouco et al., 2009). All these methodological issues may have an impact on the recovery of past GSTs from BTPs. In turn, changes in surface climate parameters, such as snow cover, evaporation or LULC over long timescales, may have an impact on the recovery of past SAT variations from reconstructed GSTs, due to the SAT–GST decoupling.

Pollack and Smerdon (2004) evaluated the impacts of some methodological issues, like the uneven geographical distribution of the borehole sites and the aggregation of the data, on hemispheric scale borehole-based reconstructions. They showed that the warming trend indicated by borehole temperature reconstructions for the NH is essentially independent of the spatial aggregation of the data and associated weighting scheme to determine the hemispheric mean. Likewise, they argued that the predominantly mid-latitude distribution of boreholes should

be able to capture the NH temperature evolution. Additional sampling issues like differences in timing and depth of borehole logs have received less attention. As most of the logs were done before 1980 CE, the aggregation of this information should be done with caution as the post 1980 years, when global warming has been larger, are underrepresented. Indeed, the methodological implementation used to obtain hemispheric to global SAT estimations from borehole-based reconstructions (e.g., Huang et al., 2000; Harris and Chapman, 2001) represent a partially muted estimation of the late industrial warming since only a small portion of BTP data includes information of the last decades of the 20th century (Pollack and Smerdon, 2004; Jones et al., 2009). Their influence is, however, difficult to assess in studies with real BTPs.

In the recent decades, uncertainties in borehole-based temperature reconstructions have been further explored with the use of climate model simulations as a surrogate reality in which the reconstruction method is implemented mimicking real-world cases in pseudo-proxy experiments (PPEs; Smerdon, 2012). The simulations can be considered as physically plausible climate realizations, compatible with the external forcings imposed and complex enough to allow for a credible implementation of the reconstruction method. The use of millennium-length GCM simulations has provided a long-term framework in which to analyze the physical background of the SAT–GST relationship. Further, they allow for the assessment of methodological issues in controlled experiments.

González-Rouco et al. (2006) tested for the first time the robustness of using the borehole method for reconstructing past SAT variations in a PPE in which the borehole methodology was replicated in the simplified reality of the ECHO-G GCM. They simulated underground temperature anomaly profiles using a one-dimensional heat-conduction forward model driven by simulated GSTs. The anomaly profiles were created at every land model grid point up to a depth of 600 m to accommodate for the temperature variations of the LM. Subsequently, they applied an inversion approach following standard procedures in borehole reconstruction strategies to reconstruct the long-term trend GST histories that could be afterward compared to the simulated SATs for verification. Their results supported the overall performance of the borehole methodology in retrieving the past SAT long–term variations. Additionally to this idealized case in which BTPs would be available for every model grid point, they also implemented the method in a more realistic case by limiting the sampling of BTPs to replicate their actual spatial distribution (Huang and Pollack, 1998). This exercise rendered similar robust results. González-Rouco et al. (2009) extended this analysis in order to include the effects of the variability in the timing of BTP logging dates and depths for a specific example over North America. Their results support estimations with the borehole technique from retrieving the North American

20th-century warming. In general, these initial experiments support the overall performance of the borehole methodology in retrieving the long-term GST variations, and additionally, they have shown that globally GST is strongly coupled to SAT over centennial timescales. Nonetheless, the lack of representation of some of the LM external forcings (i.e., LULC changes; Jansen et al., 2007) may conceal some of the effects regarding the influence of SAT–GST decoupling and affect BTP based reconstructions.

García-García et al. (2016) used the LM all-forcing simulations from state-of-the-art ESMs within the frame of the CMIP5/PMIP3 (Paleoclimate Modeling Intercomparison Project phase 3). Unlike the GCMs used in the early works that do not incorporate some external forcings, the CMIP5/PMIP3 simulations include a larger representation of LM forcings than pre-PMIP3 experiments (Fernández-Donado et al., 2013; Taylor et al., 2012), including a variety of land-surface model components. They followed the idealized approach in González-Rouco et al. (2006), sampling the full model grid over land. This allowed for demonstrating the performance of the borehole method over the current generation of ESMs, involving different land models and surface parameterizations. Their results reinforced the reliability of recovering past global surface temperature variations from subsurface temperature measurements using current state-of-the-art GCMs.

The previous assessments of the potential sources of bias on borehole climatology mostly support estimations of global/hemispheric past surface temperature histories obtained from borehole temperature inversions. First, they support the overall performance of the methodology in retrieving past GST low-frequency variations. Second, the use of up-to-date forcing representations in PMIP3/CMIP5 model ensembles under realistic scenarios involving a full setup of LM natural and anthropogenic forcings (Schmidt et al., 2011, 2012) also ensures that long-term alterations of surface properties like those induced by LULC changes, the effect of anthropogenic aerosols cooling and the potential long-term snow cover feedbacks induced by both forcings, have not been found to bias inversion results at global/hemispheric scales. More recently, Melo-Aguilar et al. (2020) have considered a more complete set of methodological constraints on a global scale, in particular, the variability in logging sites and dates of the BTPs. Likewise, all of these assessments have been done in the frame of the so-called all-forcing experiments. Melo-Aguilar et al. (2020) provided new insights into the individual effect of each of the forcings and the related feedbacks, both at global/hemispherical or at continental/regional scales that were not addressed. This helped to quantify the effect of the individual forcings on SAT-GST coupling and to obtain estimates of its particular temporal evolution that would allow for disentangling its contribution from that of other forcings.

In this Thesis, we report on the Melo-Aguilar et al. (2020) implementation of PPEs with the use of LM simulations from the CESM-LME. Chapter 5 updates analyses of methodological and physical influences on GST reconstructions and SAT–GST coupling, respectively, that had not been systematically addressed before. The PPE is designed such that the influence of methodological sampling issues at global and regional scales as well as that of individual external forcings on biases BTPs reconstructions can be addressed.

1.6 Main objectives and structure of the thesis

The land–air interactions determine the connection between SAT and GST. In turn, the conductive downward propagation of the long-term surface temperature changes provides an additional basis for the development of past temperature reconstructions from current BTPs. Understanding the mechanisms that control the interactions between the atmosphere and the ground surface, as well as subsurface heat transfer processes, is of relevance to assess potential sources of bias for the estimations of borehole-based temperature reconstructions. This is because it is widely recognized that the SAT–GST coupling may be disrupted by various mechanisms at the Earth's surface that alter the energy balance on different timescales.

Therefore, the main goal of this Thesis is to deepen our understanding of the interactions between the ground surface and the atmosphere that modulate the SAT–GST coupling assumption on various temporal and spatial scales and assess their implications for climate reconstructions based on the use of BTPs. A soil temperature observational database for Spain is developed to expand the assessments of previous works regarding both the subsurface heat transport and the SAT–GST coupling for a variety of soil types and climatological conditions. Moreover, a recent ensemble of LM simulations with the CESM provides a valuable opportunity to explore further the physical mechanisms that control the energy fluxes at the ground-atmosphere interface at different spatial and temporal scales. This Thesis also aims at updating the assessments of previous works regarding the limitations imposed by a set of methodological issues that may distort the recovery of past GSTs from BTPs. In this process, some of the conductive heat transport features of the surface temperature signals into the subsurface are also explored.

The central scope of this Thesis is achieved by the development of some specific objectives as follow:

In this Thesis, an observational soil temperature database has been created for peninsular Spain and the Balearic Islands. Chapter 2 provides a detailed description of the development of the SoTOS database that includes soil temperature records down to a depth of 1m for a total of 39 locations in peninsular Spain and the Balearic Islands. Section 2.1 describes the main characteristics of the data, the data management issues as well as the quality control (QC) procedure that was applied to ensure the homogeneity of the data. The QC process includes the detection, elimination or correction of several error typologies like spikes in the temperature records or drifts and changes in variance over segments of the time series. A final quality controlled dataset for the 1985–2018 period is presented and will be available in the future for the community.

This Thesis uses the SoTOS database to address the thermal regime in the shallow subsurface and the coupling between SAT and GST at interannual to decadal timescales over Spain. Chapter 3 employs the final quality controlled SoTOS database to evaluate characteristics of the conductive downward propagation of the surface temperature in the shallow subsurface (i.e., the upper 1 m below the surface). We analyze the rate of amplitude attenuation and phase shift with the depth of the annual temperature cycle for the different soil layers included in SoTOS. Such an analysis allows evaluating the conductive heat transport assumption as the propagation into the subsurface of surface temperature signals that can be approximated to a harmonic function of time, such as the annual cycle, implies an exponential amplitude attenuation and a linear phase shift with the depth. Chapter 3 explores these characteristics in the 39 SoTOS sites and discusses the thermal regime in the shallow subsurface, including other mechanisms that may account for the shallow soil temperature distribution. Subsequently, Chapter 3 addresses the SAT–GST coupling assumption using SoTOS. To do so, the structure of the SAT–GST covariance and the mean SAT–GST offset over the full period of data is analyzed for the annual and seasonal periods. This helps to understand the influence of different mechanisms on the SAT–GST coupling at short timescales. Finally, the SAT–GST relationship is explored on interannual to multi-decadal timescales by assessing the evolution of both variables over the 1985–2018 CE period.

The influence of external forcing factors on the long–term SAT–GST coupling over the LM is addressed in surrogate climates. Chapter 4 focuses on analyzing the SAT–GST relationship at timescales longer than a couple of decades. To do

so, LM simulations with state-of-the-art ESMs are employed as they provide a longer time frame than the instrumental records. Specifically, this chapter uses the CESM–LME that, up to date, is the largest ensemble of LM simulations with a single model, including a subset of all-forcing simulations incorporating the integrated effect of all of the LM external forcings as well as subsets of single-forcing experiments that separately consider the transient evolution of solar variability, volcanic activity, orbital changes, GHGs, anthropogenic aerosols and LULC. Therefore, the use of the CESM-LME allows for getting insights into the individual effect of each of the forcings and the related feedback on the SAT–GST relationship. The first part of Chapter 4 is centered on analyzing the structure of the SAT–GST covariance over the LM at a global scale using the all-forcing simulations. This allows assessing the relationship between SAT and GST in a realistic representation of real-world forcing conditions. This chapter discusses different cases in which the SAT–GST relationship shows deviations from the strong coupling assumption, providing a detailed description of the physical mechanisms that modulate the energy transfer at the Earth's surface, and thus, the SAT and GST responses. The second part of Chapter 4 explores the long-term relationship between the SAT an GST, focussing on the individual effect of the anthropogenic external forcings (i.e., GHGs, anthropogenic aerosols and LULC). Therefore, in this part, the GHGs, LULC and anthropogenic aerosols single-forcing experiments are considered to separate each forcing's effect in detection and attribution experiments (Hegerl et al., 2011). This chapter's results show that the different mechanisms may disrupt the long-term SAT–GST coupling at local to regional scales. Among them, long-term snow cover feedbacks due to the effect of GHGs in addition to LULC changes that alter the energy balance at the surface stand as the main contributors to long-term SAT–GST decoupling. These long-term SAT–GST decoupling processes may constitute a source of bias for borehole temperature reconstructions since the downward propagating thermal signal may be uncorrelated to long-term SAT variations.

Addressing the potential sources of bias in borehole temperature reconstructions considering both methodological aspects and the physical SAT–GST decoupling in a pseudo-proxy framework. Chapter 5 is built on the findings of Chapter 4 that shows that long-term SAT–GST decoupling may exist due to the influence of anthropogenic external forcings. Chapter 5 also uses simulations from the CESM-LME in to follow up on Chapter 4 results. This part of the Thesis develops a PPE strategy in which a borehole reconstruction strategy is implemented in the world of the CESM-LME to assess the technique's performance for a range of spatial scales. The PPE consists of the simulation of synthetic anomaly temperature profiles using a heat-conduction forward model driven by simulated surface

temperature. Subsequently, the anomaly temperature profiles are inverted using singular value decomposition (Mareschal and Beltrami, 1992) to retrieve the forcing surface temperature signal. The PPE is designed to evaluate the influence of sampling methodological issues and physical SAT–GST responses at global and regional scales by considering the actual distribution of BTPs and their depths and logging times. Thus, mimicking the real-world scenario of the borehole reconstruction technique. Chapter 5 provides an updated analysis of methodological constraints on GST reconstructions and, most importantly, physical influences on SAT–GST coupling that had not been systematically addressed before. Further, this study evaluates for the first time the individual effects of the GHG and LULC external forcings on the performance of the borehole method to retrieve past SAT changes. Results indicate that the influence of external forcings and sampling issues can be severe at regional scales.

2

Data: observations and model simulations

This chapter describes the sets of data employed in this Thesis, both instrumental records and modeling sources. First, Section 2.1 is centered on the process of the development of the SoTOS observational database. This includes the compilation process, a general description of the available data, and the QC procedure applied to assure their quality. Second, Section 2.2 provides a brief description of the ERA5-Land dataset (ERA5-L; Muñoz Sabater, 2019), which is a reanalysis product developed by the European Centre for Medium-Range Weather Forecasts (ECMWF), that combines land-surface modeling with observations to produce fields of the evolution of the land surface states and fluxes from 1981 to present. The SoTOS and ERA5-L datasets are used to develop the analyses of Chapter 3. Finally, the basic characteristics of the ESM ensemble simulations from the CESM-LME analyzed in the present study (Chapters 4 and 5) are described in Section 2.3.

2.1 Soil temperature observational dataset for Spain (SoTOS)

A set of soil temperature (ST) observations were collected from meteorological stations from the Spanish Meteorological Agency (Agencia Estatal de Meteología; AEMET) for a total of 39 sites with relatively good coverage of most of peninsular Spain and the Balearic Islands. The ST data is measured at 0.05, 0.10, 0.20

0.50 and 1 m depth (ST_{005}, ST_{010}, ST_{020}, ST_{050} and ST_{100}, hereafter). Besides, 2m air temperature (SAT) has also been collected from the same 39 sites. The SAT and ST_{005}, ST_{010}, ST_{020} data are provided with 6-hourly resolution with measurements available at 00:00, 07:00, 13:00 and 18:00 GMT; for SAT, daily max and min values are also available. For the ST_{050} and ST_{100}, only one observation per day is taken at 18:00 GMT. For the development of SoTOS, data spanning from 1985 to 2018 have been employed.

Table 2.1 provides a general description of the meteorological observatories included in the dataset. Figure 2.1 illustrates the spatial distribution of the 39 sites and information regarding the initial and final date of available records at each of the sites. Additionally, the total length and percentage of missing values in the time series are shown in Figure 2.1 (right). The information is displayed as an example for SAT, and the ST_{005} and ST_{100} levels.

The maps in Figure 2.1 are constructed from daily averages of data that have been calculated from the available intradaily records. For the SAT, the daily mean is obtained as the average between the daily maximum and minimum values since both of these records are available. This is a standard procedure to obtain daily average in meteorological temperature data (WMO, 2018) and is also the way daily means are calculated by AEMET (personal communication). For ST_{005}, ST_{010} and ST_{020} daily maximum and minimum values are not available. Therefore, the daily mean is calculated using the minimum and maximum values among the 4 available daily measures. The minimum temperature value is regularly captured at 07:00 GMT for soil layers close to the surface, whereas the maximum varies between the records at the 13:00 and 18:00 depending on both depth and date of the year. The latter results from the downward propagation of surface temperature variations, such as the daily cycle, which is characterized by a linear phase shift and exponential amplitude attenuation with the depth due to the nature of heat conduction (see Section 1.3).

The SAT data present a relatively homogeneous distribution of initial and final dates of available records in 1985 and 2018, respectively, with only a few cases where the records start later or end at an earlier date (i.e., 4, 26 and 39). On the other hand, the length of available ST data is highly variable among sites as well as for the different available soil layers due to variability in the initial and ending dates for ST records (Figure 2.1d,e,g,h). This situation implies that the ST data do not cover the full 1985-2018 period at all sites. In addition, missing data also affect the availability of records (Figure 2.1c,f,i). Note the variability in the circle size and color for both ST_{005} and ST_{100} indicating a variable length of available data and gaps of missing information, with a few cases in which the missing data can reach up to 60-80% (e.g., Figure 2.1f; site-6).

Table 2.1: List of AEMET weather stations included in SoTOS.

Site	Code	Altitude (m)	Longitude (^{o})	Latitude (^{o})	Location
1	4642E	19	353.0903	37.2800	Huelva (Ronda este)
2	7178I	62	358.8306	38.0028	Murcia
3	4452	185	353.1708	38.8833	Badajoz (Talavera, base aérea)
4	4116I	622	356.2597	38.9508	Almagro (Famet)
5	4121	627	356.0803	38.9894	Ciudad Real
6	8178D	674	358.1392	39.0069	Albacete (Observatorio)
7	3469A	405	353.6606	39.4722	Cáceres (Carretera Trujillo)
8	B691	40	3.0197	39.7464	La Puebla
9	3260B	516	355.9506	39.8847	Toledo (Buenavista)
10	8096	956	357.8619	40.0667	Cuenca
11	8368U	900	358.8772	40.3517	Teruel
12	2444	1130	355.3214	40.6500	Ávila (Observatorio)
13	3168C	635	356.8258	40.6611	Guadalajara (El Serranillo)
14	2462	1890	355.9897	40.7806	Navacerrada (Puerto)
15	2465	1005	355.8728	40.9478	Segovia (Observatorio)
16	2867	790	354.5039	40.9456	Salamanca (Matacán)
17	0200E	412	2.1253	41.4194	Barcelona (Fabra)
18	2614	656	354.2664	41.5167	Zamora (Observatorio)
19	0222	168	2.1689	41.6139	Caldes de Montbui
20	9771C	192	0.5950	41.6258	Lleida (Observatorio)
21	2422	735	355.2333	41.6500	Valladolid (Observatorio)
22	9434	247	358.9919	41.6619	Zaragoza (Aeropuerto)
23	2539	846	355.1500	41.7000	Valladolid (Villanubla)
24	0149D	280	1.8392	41.7211	Manresa (La Culla)
25	2030	1082	357.5333	41.7667	Soria (Observatorio)
26	0341	611	2.2278	41.8478	Tona (Escola)
27	1495	255	351.3680	42.2236	Vigo (Peinador)
28	1690A	143	352.1397	42.3278	Ourense (Granxa Deputación)
29	2331	890	356.3675	42.3561	Burgos (Villafría)
30	1484C	107	351.3836	42.4400	Pontevedra (Mourente)
31	1549	534	353.4000	42.5639	Ponferrada
32	2661	916	354.3506	42.5889	León (Virgen del camino)
33	9091O	508	357.2772	42.8839	Vitoria
34	1428	364	351.5731	42.8994	Santiago de Compostela
35	1505	444	352.5439	43.1153	Rozas (Aerodromo)
36	1024E	252	357.9605	43.3075	San Sebastián (Igueldo)
37	1249I	336	354.1267	43.3536	Oviedo (El Cristo)
38	1387	58	351.5806	43.3672	A Coruña
39	1111	52	356.2003	43.4917	Santander

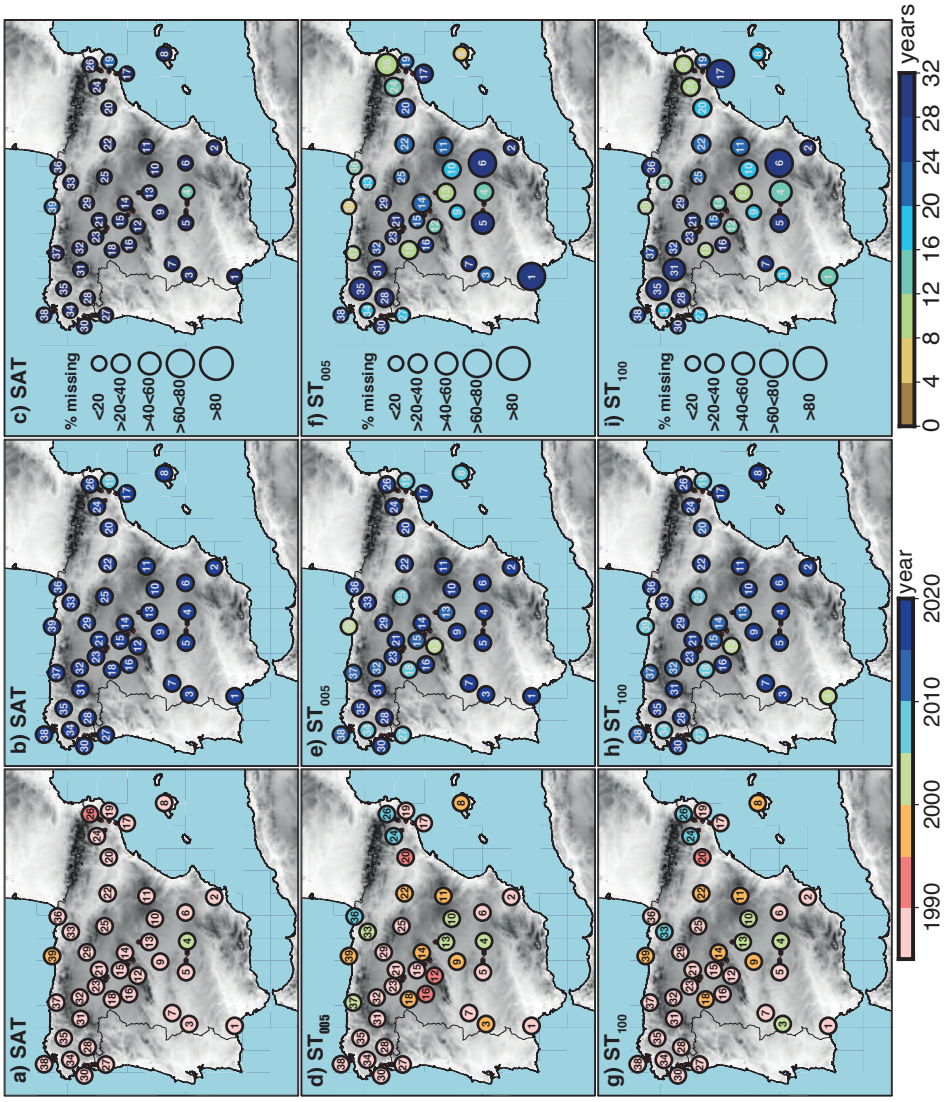

Figure 2.1: Initial (left) and final (middle) date of records for each of the 39 sites considered in SoTOS. Percentage of missing values relative to the time interval between the initial final dates of observations (right). From top to bottom, surface air temperature (SAT) and soil temperature (ST) at 5 and 100 cm are shown as example. The arrows at some sites (e.g., 4, 5) point to the actual location while the symbol has been moved to avoid overlap.

2.1.1 Quality control

Measurements of meteorological or climatological data are subject to errors that can occur at different stages during the process of data manipulation, starting with the recording, storage and other non-climatic issues that may lead to systematic or uncontrolled errors in the final records (Lucio-Eceiza et al., 2018a,b). This situation implies that some QC is required previous to using the data in any meteorological or climatological application. According to this, quality control of the original 6-hourly data and the SAT daily maximum and minimum values have been performed. The QC consists of a three-step process that identifies suspicious data that are flagged as measurement errors. The undoubtedly wrong data is discarded, whereas some other issues have been corrected according to standard procedures. The following sections describe the quality control implemented in this study.

2.1.1.1 Internal consistency check

The first step consists of detecting extreme values that are wrong with a very high level of confidence. Such errors in the datasets may be related to problems in the digitalization process, typos or different values in reporting data (Hu and Feng, 2003; Lucio-Eceiza et al., 2018a,b). Some of these issues are easy to identify from a simple visual inspection. However, it is essential to distinguish between data outliers due to extreme meteorological events from those values that are undoubtedly measurement errors. In order to do so, we have established two basic criteria: (i) unrealistically large or small temperature values based on absolute values above 50°C and below -50 °C following Gonzalez-Hidalgo et al. (2015) and (ii) spikes in temperature data (Reek et al., 1992). The latter has been identified by taking the difference between three consecutive measurements, centered on each of the data points. The smallest absolute difference is considered and compared with a threshold value. If the difference exceeds this threshold, the data point is flagged (spike) and deleted. For the second criteria, imposing a correct limit to determine whether a certain data point is actually an error is crucial to avoid eliminating extreme values that are due to meteorological extremes. Thus, three different threshold values have been compared. Firstly, this limit is established based on the standard deviation (s) and the mean ($\overline{x}$) of the smallest-absolute-difference set for the complete time series. For instance, a first limit value is imposed at $8 \times s$. If the smallest-absolute-difference of any data point is greater than $8 \times s$, that data point is flagged and deleted (Hu and Feng, 2003). A second threshold was imposed at $10 \times \overline{x}$, and the same comparison is

made to flagged the suspicious data. Secondly, a third threshold is imposed at differences greater than a reference value of 10 °C (Meek and Hatfield, 1994). Selecting any of these thresholds yield similar results. However, the use of either the $8 \times s$ or the $10 \times \overline{x}$ thresholds imposes a relatively low limit. This results in flagging some data points that are not clearly "spikes". Rather, they may be associated with actual meteorological extremes. Therefore, the 10 °C limit is selected as the reference threshold since it represents a more conservative approach. Figure 2.2a illustrates results of the internal consistency check for site-16. The wrong data are flagged (red points) and deleted. A summary of the total number of flagged data points for each of the 39 stations and all of the ST layers is shown in Figure 2.3.

2.1.1.2 Changes in long term trends

Other situations affect the quality of temperature data, such as changes in the average or trends over some periods of the time series, i.e., inhomogeneities. In the case of soil temperature data, these types of changes can be identified from a visual inspection of the data given the smooth character of the time series at depth and the limited number of sites. Plotting the ST data for the different available layers allows comparing their evolution and identifying whether a certain time series is affected by this situation. Hu and Feng (2003) showed that changes in the average/trend in some of the ST layers can be corrected in some cases. The latter is possible if this behavior is present over short periods of time, and additionally, if there is available information from either a different soil layer for the same station that is not affected by this type of problem or from a neighboring weather station that can be used as a reference. Similar to the correction of inhomogeneities in other meteorological parameters (González-Rouco et al., 2001).

Figure 2.2b illustrates a case in which the original data presents a change in the trend over a short period of time (shaded region; 1990–1994) for one of the soil layers (ST_{020}) as indicated by the red line. Such a change in the trend is not evident at the other layers (ST_{010} is shown for comparison). In these cases, the estimation of the corrected data for the affected segments can be done by using a reference time series. This is possible because the soil temperature at different soil layers is a function of the downward propagation of the surface temperature signal through conductive heat transport. Although the conductive propagation implies an exponential amplitude attenuation and linear phase shift with the depth relative to the surface temperature signal, the ST at different soil layers contains a similar long-term behavior (trend) provided that shallow depths are

considered. Thus, the reference depth trend can be used to estimate the corrected data for the segment in the affected one. With this in mind, Hu and Feng (2003) constructed the following equation:

$$ST_e = ST_c - (a_c + b_c t) + (a_r + b_r t) - \Delta ST_{rc} \tag{2.1}$$

Here, ST_e and ST_c represent the estimated and the original ST in the affected segment, respectively, t is the time in days, and $a_c + b_c t$ and $a_r + b_r t$ represent the linear fit to the data for the affected and the reference ST in the affected segment, with a and b as the regression coefficients. ΔST_{rc} stands as the difference in the temperature average between the affected and reference ST, excluding the affected segment; it is included to account for the temperature variations with depth.

Equation (2.1) is employed to estimate the corrected data for the segment of the ST_{020} in the example depicted in Figure 2.2b (dark-grey line). In this case, ST_{010} is used as the reference temperature. This method is employed for sites 1, 5, 8, 23, 27 and 36. A summary of the corrected segments and the percentage of fixed records relative to the total length of the series is shown in Table 2.2.

2.1.1.3 Changes in amplitude

Similar to the changes in the average of time series, there are cases in which there are changes in the amplitude over segments of time series that further affect the quality of the data. This type of error may be related to changes in either the location or instrumentations of weather stations or due to difference in the reporting units among others. There are several methods to detect a change in amplitude in time series that goes from cumulative sum of squares to Bayesian techniques (Killick et al., 2010). We used the scale cumulative sum (SCUSUM), which is a variant on the cumulative sum developed to detecting changes in the amplitude of time series (Peterson et al., 1998). Variations in the amplitude in the observations would appear as changes in the slope of the SCUSUM plot (Rhoades and Salinger, 1993); segments of high (low) SCUSUM indicate that the variance is larger (smaller) than the variance of the whole time series.

A test statistic (Page, 1955) allows for a quantitative evaluation of the degree of change. For a series of observations $X_1, X_2, ..., X_n$, the SCUSUM test statistic is defined as:

$$\max_{i \leq n}(S_i - \min_{j \leq i} S_j), \tag{2.2}$$

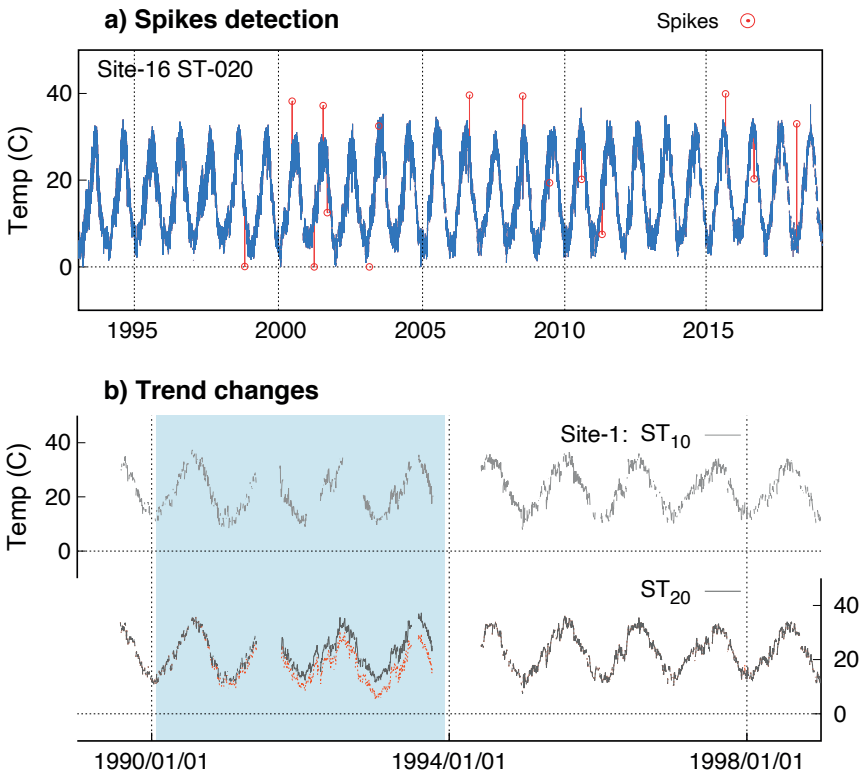

Figure 2.2: Example of the spikes detection for the soil temperature at 0.2 m depth in site-16 (a) and trend changes in time series (b) for site-1 as an example. Red circles in a) stand for erroneous flagged and corrected data. The shaded area in b) highlights the period (1989-1994) in which the ST at 0.2 m depth presents a change in the trend that is not evident in the ST at the other available layers. In this case, ST at 0.1 m depth is used as a reference for the correction of the change in the ST at 0.2 m depth.

with S_i as the SCUSUM score:

$$S_i = \sum_{j=1}^{i} [(X_j - \overline{X})^2 - s^2], \tag{2.3}$$

where j, i and n are the first, the i-th data point and the total number of data points, respectively, $\overline{X}$ is the sample mean and s^2 is the variance.

The SCUSUM test statistic provides a measure of the magnitude of the change in the variance for a given time series (Peterson et al., 1998). It allows for testing the statistical significance of the change. Values larger than a threshold indicate significant changes in the variance. Once the sites with significant changes in variance are identified ($p < 0.05$), the SCUSUM scores graph is evaluated to identify the time series affected segments, enclosed by the range of minimum and maximum SCUSUM scores.

We applied this procedure to the 39 sites to both the SAT and the ST for the 5 available layers. The SCUSUM test statistic analysis and the SCUSUM graphs indicate that the STs at all depths for site-29 present a significant change in the variance over the time interval 2009–2018.

In some cases, it is possible to obtain corrected ST data for the segments that exhibit such type of error (Hu and Feng, 2003). To do so, a reference temperature series not affected by such a change in the amplitude is required. The reference temperature series may be the ST at a different depth of the same station or from a neighboring station if temperature variations are comparable. Additionally, the reference series should be of similar amplitude and similar average over the non-affected segments. Once these requirements are fulfilled, the following equation is provided by Hu and Feng (2003) to estimate the corrected data for the affected segment:

$$ST_e = [(ST_c - \overline{ST}_c)/s_c](s_r s_{c1}/s_{r1}) + \overline{ST}_c - \Delta ST_{rc}, \tag{2.4}$$

where ST_e and (ST_c) represent the estimated (affected) ST, respectively, $\overline{ST}_c$ is the average of the affected ST, s_c and s_r are the standard deviations of the affected and the reference series, respectively, for the erroneous data segment; s_{c1} and s_{r1} are standard deviations of the affected and the reference temperature series excluding the erroneous data segment. The average soil temperatures difference between affected and reference series, excluding the erroneous data segment, ΔST_{rc}, is included in Equation 2.4 to account for soil temperature variation with depth.

Equation 2.4 is employed to estimate the temperature with the correct amplitude for only one of the sites included in this work, i.e., site-29, in which the original data is affected by a change in amplitude over a segment of the time series spanning from 2002 to 2008 for the ST_{010}, ST_{020}, ST_{050} and ST_{100} (not shown).

2.1.2 Final quality controlled daily surface air and soil temperature dataset

After applying the QC procedure to the original data, the erroneous measurements have been removed or corrected. The result is a quality-controlled soil temperature dataset spanning from 1985 to 2018. It is worth noting that there is some variability in the temporal coverage among the 39 sites considered herein. In addition, there are some cases in which the available data for the different soil layers included in SoTOS also presents differences in the temporal coverage.

SoTOS is employed to analyze the thermal regime in the shallow subsurface, and the SAT–GST coupling over peninsular Spain and the Balearic Islands presented in Chapter 3. In spite of the variability in the temporal coverage among the 39 sites and the length of records for the different soil layers, the available information is sufficient to ensure the quality of the assessments regarding the subsurface thermal regime developed in Chapter 3. Also, for many sites, the length of the ST_{005} and SAT data permits the assessment of the SAT and GST relationship at inter-annual to decadal timescales.

A summary of the quality-controlled data is presented in Figure 2.3 and Table 2.2. In general, the flagged and deleted data from the internal consistency check was less than 1% of the data for each of the 39 sites for both SAT and ST at all depths. Corrections related to changes in the trend or variance over some segments of the series were implemented over a total of 6 and 1 sites, respectively. All of these corrections were applied to ST data at different depths. As these segments were adjusted, the estimated corrected data is included in SoTOS.

Table 2.2: Summary of the corrected ST data according to the changes in trend or amplitude over some segments of the time series, indicating the site number (Table 2.1), the ST layer, the time interval of corrected records and the number of records corrected.

	Site	ST layer	Corrected period	% of data
Changes in trend	1	ST_{020}	1990/08/31-1994/02/28	20.8
	1	ST_{100}	1992/08/04-1997/10/31	55.4
	5	ST_{100}	1987/01/01-1988/12/31	8.8
	8	ST_{100}	1998/07/01-2003/02/28	30.9
	23	ST_{100}	2005/01/01-2007/12/31	9.3
	27	ST_{100}	1989/01/01-1995/12/31	38.7
	36	ST_{020}	1995/01/01-1999/03/31	13.2
Changes in amplitude	29	ST_{020}-ST_{100}	2002/01/01-2008/01/31	29

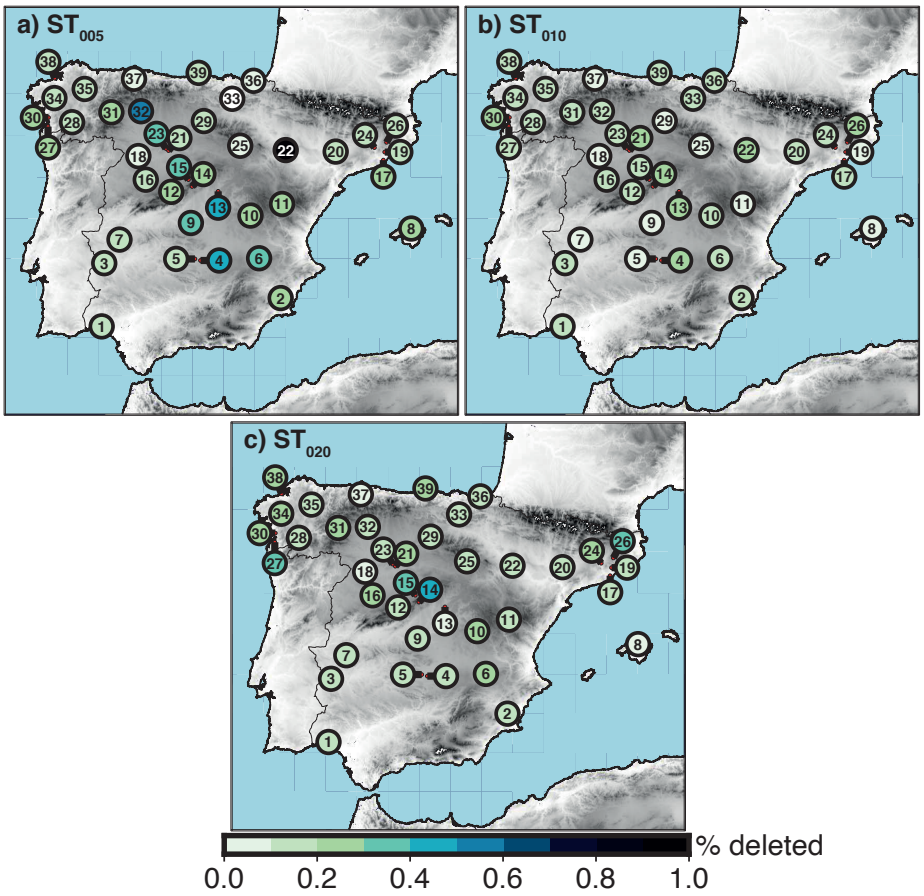

Figure 2.3: Percentage of flagged and deleted data relative to the total of observations for the 39 sites in the SoTOS for ST_{005} (a), ST_{010} (b) and ST_{020} (c). The maps for SAT, ST_{050} and ST_{100} have percentages of deleted data that do not exceed 0.1% (not shown).

2.2 ERA5-Land reanalysis

The present work considers data from the ERA5-L dataset, a reanalysis approach that combines land-surface modeling with observations to produce fields of the evolution of the land surface states and fluxes. ERA5-L is forced by atmospheric variables such as air temperature and air humidity from the ERA5 climate reanalysis (Hersbach et al., 2020). ERA5-L runs at a spatial resolution of 9 km, representing an enhancement relative to the ERA5 climate reanalysis resolution (31 km). The temporal frequency of the output is hourly covering the period from 1981 to present and the fields are masked for all oceans. ERA5-L runs the stand-alone land surface model, i.e., not coupled to the atmospheric or oceanic components, of the ECMWF Integrated Forecasting System (IFS). Currently, ERA5-L contains only one high resolution realization.

The ERA5-L provides soil temperature and other soil variables at 4 different levels as defined by the H-TESSEL land surface model (IFS, 2018) within the ERA5-L reanalysis. The H-TESSEL is discretized in 4 layers (Table 2.3) with a bottom boundary condition placement (BBCP) of no heat flux of energy at 2.89 m. The soil temperatures are defined at full layers while the heat fluxes are defined at the interface between layer i and $i+1$. Although the vertical distribution of soil levels in the H-TESSEL land surface model shows a relatively good agreement to the soil layers included in SoTOS, the model soil layer depths do not exactly match those of the observations. Therefore, the model soil temperatures are linearly interpolated between nodes of each soil layer to match the SoTOS soil layers (i.e., 0.05, 0.1, 0.2, 0.5 and 1 m depth), thus allowing for contrasting the state of the soil thermodynamics in observational and reanalysis data in a comparable framework. For some other analyses regarding the soil thermal state, the model soil temperatures and volumetric water content are considered in their original vertical resolution (Table 2.3); further details are provided in Chapter 3.

Specifically, soil temperature data from ERA5-L over the region between 35–45° N and 10° W–5° E, including peninsular Spain and the Balearic Islands, for the 1985-2018 period at monthly temporal resolution is employed to provide a comparison between SoTOS and the soil temperature fields as represented by state-of-the-art land surface modeling products. Additionally, other land variables from ERA5-L, such as volumetric soil water content and 2m air temperature, are also considered for some specific analyses developed in Chapter 3 since they are of interest for analyzing the soil thermal state and assessments of the land–air interactions, thus allowing for describing the overall representation of the soil thermal state in observational records and reanalysis information.

Table 2.3: Soil layers and node depths in the H-TESSEL land surface model. The node depth, the depth at which the soil temperatures are defined, corresponds to the middle point at each layers.

Layer	**Layer depth (m)**	**Node depth (m)**
L1	0–0.07	0.035
L2	0.07–0.28	0.175
L3	0.28–1	0.64
L4	1–2.89	1.945

2.3 Community Earth System Model–Last Millennium Ensemble (CESM-LME)

In this Thesis, output from simulations with the CESM-LME spanning from 850 to 2005 CE are employed to develop the analyses described in Chapters 4 and 5 regarding the long–term coupling between SAT and GST as well as an overall assessment of the potential sources of bias in borehole temperature reconstructions from the methodological constrains and some physical mechanisms. These CESM-LME simulations were produced with version 1.1 of the CESM (CESM1; Hurrell et al., 2013). The Community Atmosphere Model version 5 (Neale et al., 2012) is used as the atmospheric component and the Parallel Ocean Program version 2 (Smith et al., 2010) represents the ocean component, including the Los Alamos sea ice model (Hunke et al., 2015). The horizontal resolution of the CESM-LME is $\sim 2^{\circ}$ over the atmosphere and land and $\sim 1^{\circ}$ in the ocean and sea ice (Otto-Bliesner et al., 2016).

The land surface component in the CESM1 is the Community Land Model Version 4 (CLM4; Lawrence et al., 2011), which incorporates some improvements relative to the previous version of the same model (CLM 3.5; Oleson et al., 2008) in the representation of land-surface processes that are important for the energy transfer between the atmosphere and the soil (Oleson et al., 2010). Some of these include a better description of ground evaporation, thermal and hydrologic properties of organic soil, the ground depth, snow albedo, snow cover fraction and burial fraction of vegetation by snow (Lawrence et al., 2011). In addition, the CLM4 has the deepest BBCP among the current generation of land surface models to date (Cuesta-Valero et al., 2016), placed at 42.1 m depth, discretized into fifteen layers (see Table 2.4) and including up to five additional layers in the overlying snowpack. This is relevant for borehole reconstruction applications because

Table 2.4: Soil layers and node depths in the CLM4. Note that the node depth, that are the depth at which the thermal properties are defined for soil layers (Oleson et al., 2010), does not necessarily coincide with the center of the layer.

Layer	**Layer depth (m)**	**Node depth (m)**
L1	0.017	0.007
L2	0.045	0.027
L3	0.090	0.062
L4	0.165	0.118
L5	0.289	0.212
L6	0.492	0.366
L7	0.828	0.619
L8	1.382	1.038
L9	2.296	1.727
L10	3.801	2.864
L11	6.284	4.739
L12	10.377	7.829
L13	17.125	12.925
L14	28.252	21.326
L15	42.103	35.177

previous works have shown that shallow BBCPs may not represent reliably the downward propagation of the temperature signal corrupting the amplitude attenuation and the phase shift (Smerdon and Stieglitz, 2006; Nicolsky et al., 2007; Alexeev et al., 2007). Therefore, this improved and deeper land surface model allows for analyzing the coupling between SAT and GST with a more realistic atmosphere-subsurface heat transfer scheme.

Simulated soil temperature at different depths are employed to describe some of the characteristics of the soil thermal state as represented by long integrations within the CESM-LME in both Chapter 4 and 5. The subscript notation ST_{L1}, ST_{L2}, ..., ST_{L15} is used to describe the soil temperature layers along the document. Other variables are also employed to develop the analyses regarding the land–atmosphere interactions: 2m air temperature, fraction of ground covered by snow, soil water content in the top 10 cm of soil, surface latent and sensible heat

fluxes, net solar radiation at the surface, reflected shortwave radiation, soil ice content, etc.

Table 2.5: Simulations and LM external forcing reconstructions used in CESM-LME. The single forcing simulations span from 850 to 2005 CE, except those of anthropogenic ozone and aerosols that span the period 1850 to 2005 CE. Legend for external forcing: SOL, changes in total solar irradiance; VOLC, volcanic activity; GHG, concentrations of the well-mixed greenhouse gases CO_2, CH_4, and N_2O; LULC, land use land cover changes; ORB, orbital variations; and OZ/AER, anthropogenic ozone and aerosols. The ALL-Fsimulations contain all the single-F reconstructions specified above. Note: The current number of available simulations may be larger since additional ensemble members with the same or different forcing reconstructions are still being developed.

Forcing	**No. of simulations**	**Reference**
ALL-F	13	-
SOL	4	Vieira et al. (2011)
VOLC	5	Gao et al. (2008a)
GHG	3	MacFarling Meure et al. (2006)
LULC	3	Pongratz et al. (2008) dataset, spliced to Hurtt et al. (2009) at 1500 CE. The only plant functional types (PFTs) that are changed are those for crops and pasture; all other PFTs remain at their 1850 control prescriptions.
ORB	3	Yearly orbital position (eccentricity, obliquity and precession) is adjusted following Berger et al. (1993).
OZ/AER	2	Fixed at the 1850 control simulation values until 1850 and then include the evolving anthropogenic changes to 2005. Stratospheric aerosols are prescribed in the model as a fixed single-size distribution in the three layers in the lower stratosphere above the tropopause. The ozone forcing is from the Whole Atmosphere Community Climate Model (WACCM).

The CESM-LME incorporates simulations forced by the complete set of agreed CMIP5 LM natural (solar, volcanic and orbital) and anthropogenic (GHG, anthropogenic aerosols and LULC changes; Schmidt et al., 2011, 2012), the so–called "all-forcing" experiments (ALL-F hereafter) that were chosen following those used in Landrum et al. (2013). In addition, simulations forced by single–forcing factors are also included (Single-F hereafter). The CESM-LME is composed by more than 30 simulations including a subset of 13 ALL-F simulations and smaller subsets of Single-F simulations (see Table 2.5 for details and references therein). The ensemble of ALL-F simulations as well as some of the Single-F subensembles of simulations have been employed herein. The simulations used in this study and how they trace back to the original experiment names published by Otto-Bliesner et al. (2016) are shown in Table 2.6.

Table 2.6: Simulations used in this study from the CESM-LME. The first and second columns present the Acronyms used in this manuscript for ensembles and ensembles members respectively. The id of the original experiment files is provided in column 3.

Ensemble acronym	**Ensemble member**	**Simulation id**
ALL-F	ALL-F_i	b.e11.BLMTRC5CN.f19_g16.0i
	i=1,...,13	i=01,...,13
GHG–only	GHG_i	b.e11.BLMTRC5CN.f19_g16.GHG.00i
	i=1,2,3	i=1,2,3
LULC–only	$LULC_i$	b.e11.BLMTRC5CN.f19_g16.LULC_HurttPogratz.00i
	i=1,2,3	i=1,2,3
OZ/AER–only	OZ/AER_i	b.e11.BLMTRC5CN.f19_g16.OZONE_AER.00i
	i=1,2	i=1,2

2.4 Conclusions

This chapter describes the data sets that will be considered for this work, including instrumental records and output from modeling assessments. First, the

development of a soil temperature observational dataset for peninsular Spain and the Balearic Islands is described. This process includes the compilation of the instrumental records and the subsequent QC procedure focusing on flagging, elimination and correction of erroneous measurements. The ST information at 5 different soil layers (0.05, 0.10, 0.20, 0.50 and 1 m depth) was collected from 39 meteorological stations for Spain. Besides, SAT records for the same 39 sites were also obtained. All of the data was provided by the Spanish Meteorological Agency (AEMET). For this version of the dataset, the 39 sites, with a total of $\sim 5 \times 10^6$ records, were subject to the QC in which 6193 records were flagged and deleted as erroneous measurements ($\sim 0.19\%$) while 23817 ($\sim 0.47\%$) records were corrected either from changes in trends or changes in variance over segments of the series (see Table 2.2). Therefore, the amount of quality control corrections is relatively small. It is worth noting that the final quality controlled dataset has improved quality but is not necessarily error free as future versions may incorporate additional test not included here. The final quality controlled dataset is presented at a daily temporal resolution for the 39 sites covering the 1985–2018 period. This dataset will be made freely available to interested users after publication of the analysis presented in Chapter 3.

Second, other data sources from modeling assessments that are also considered in specific points of Chapters 3 are presented. Particularly, the ERA5-L, which is a model-based approach that uses advanced land-surface modeling to produce high-quality fields of land surface states and fluxes. Some of the output from ERA5-L is employed in Chapters 3 to compare the evolution with time to the results obtained from the analysis developed with SoTOS, assuming that ERA5-L provides a physically consistent set of land variables.

Finally, the main features of the ensemble of simulations from the CESM-LME that is considered in Chapters 4 and 5 are summarized. This includes a reference to the model components (i.e., atmosphere, ocean and land-surface), particularly a brief description of some of the characteristics of the CLM4, which is the land surface component of the CESM model. The latter incorporates the deepest BBCP among the current generation of land surface model discretized into fifteen layers (Table 2.4). This feature makes the CLM4 especially suited for the analyses developed in this thesis since deeper BBCP allows for a better representation of the propagation with the depth of the surface temperature signal compared to a shallow BBCP (Alexeev et al., 2007). Similarly, a description of the LM external forcings and simulations considered in the CESM-LME has been provided, with a focus on the All-F ensemble of simulations, and specifically, on the Single-F sub-ensemble of simulations that are employed in Chapters 4 and 5.

3

Shallow soil thermal regime and SAT–GST coupling: a case study over Spain

This chapter explores the shallow subsurface (i.e., 1 m depth) thermal regime and the SAT–GST relationship over Spain. For this purpose, the soil temperature observational dataset for Spain (SoTOS; see Chapter 2 for details) that spans from 1985 to 2018 period is employed. First, Section 3.2.2 addresses the characteristics of the conductive heat transport of surface temperature variations into the subsurface, i.e., exponential amplitude attenuation and linear phase shift with the depth, in the 39 sites included in SoTOS. The location of each site presents various land surface characteristics, soil physical properties, and climatic conditions, thus exploring conductive heat transport under different conditions. Such analysis expands applications developed in previous studies regarding the propagation of surface temperature signals in the shallow subsurface (e.g., Smerdon et al., 2003, 2004). Subsequently, an assessment of the SAT–GST covariance structure and the analysis of the evolution of the SAT–GST relationship at interannual to decadal timescales is presented in Section 3.2.3. Finally, Section 3.3 summarizes the main results of this chapter. The results obtained from the SoTOS database are compared to the ERA5-L reanalysis to assess the consistency between observational and reanalysis products. These model-data comparison exercises provide valuable information regarding the performance of reanalysis datasets, which is of interest because these products are employed for several scientific purposes and practical applications such as weather prediction, drought and flood risk management and soil moisture feedbacks on climate (Gallego-Elvira et al., 2016).

3.1 Methods

The propagation of the surface temperature variations into the subsurface is controlled by the one-dimensional time–dependent heat conduction equation (Carslaw and Jaeger, 1959):

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2}, \tag{3.1}$$

with κ as the thermal diffusivity of the medium while z and t stand for depth and time, respectively. Assuming that the temperature variations at the surface are represented by a simple harmonic signal and a homogeneous subsurface with no internal heat production, the solution of the steady-state analytic solution of the one-dimensional heat conduction equation as a function of time and depth takes the form:

$$T(z,t) = Ae^{(-z\sqrt{\pi/\tau\kappa})} cos\left(\frac{2\pi}{\tau}t - z\sqrt{\pi/\tau\kappa}\right), \tag{3.2}$$

where A and τ represent the amplitude of the temperature change and the period of oscillation, respectively. Equation 3.2 shows that the amplitude of the temperature variations is attenuated exponentially with depth. Likewise, the argument of the trigonometric factor indicates that there is a linear phase shift with depth relative to the initial phase of the oscillation at the surface ($\phi = z\sqrt{\pi/\tau\kappa}$). The amount of amplitude attenuation and phase shift are dependent on the depth, the period of the surface sinusoid, and the thermal properties of the subsurface.

Herein, the characteristics of the conductive propagation of the surface temperature into the subsurface, i.e., the exponential amplitude attenuation and linear phase shift with depth, are evaluated in the 39 sites included in SoTOS. This is achieved by extracting the amplitude and phase of the annual signal from the daily soil temperature measurements at 0.05, 0.1, 0.2, 0.5, and 1 m depth using linear least-squares fitting of the data to a sine-wave model of the form:

$$T(t) = Asin\left(\frac{2\pi}{\tau}t - \phi\right) \tag{3.3}$$

The annual signal is employed as a reference period of oscillation because the seasonal cycle is closely approximated to a simple-harmonic wave. Likewise, the annual signal is traceable within the upper meter of the soil, whereas higher frequency oscillations (e.g., daily cycle) are completely attenuated within the first meter (Putnam and Chapman, 1996). The first soil layer (ST_{005}) is used as

a reference to estimate the phase shift of the subsequent layers (i.e., ST_{010}, ST_{020}, ST_{050}, and ST_{100}). Once the amplitude and phase shift of the annual signal have been obtained for all of the available soil layers, a linear regression of both the natural logarithm of the amplitude and phase shift vs. depth is performed. In either case, the regression slope represent the wave vector ($k = \sqrt{\pi/\tau\kappa}$) of the downward propagating thermal wave (Smerdon et al., 2003). Knowing the wave vector, it is possible to estimate the apparent thermal diffusivity as an average over the entire soil profile and throughout the year. This can be seen as an estimation of the conductive propagation of the surface temperature signal in the shallow subsurface. This procedure is applied to the 39 sites included in SoTOS in order to provide an estimation of the variability of the regional thermal diffusivity in Spain. Therefore, this analysis permits testing the subsurface conductive heat transport assumption for various sites that include different climatological and thermal properties of the soil, thus expanding on the analysis of previous works (e.g., Smerdon et al., 2004).

For a correct implementation of this procedure, it is convenient that temperature records for all soil layers at individual locations have the same time span. However, due to the variability in the temporal coverage of available records for the different sites included in SoTOS (see Chapter 2), there is not a common period of time in which all of the sites include continuous records at all soil layers. Further, there is also some variability in the available data for the different soil layers at individual locations in some cases. Despite this, all SoTOS sites contain enough information to ensure that soil temperature records are available at all soil layers during a common time period, although the specific period of data availability may differ slightly from site to site. A condition of a minimum of 5 years of continuous data is imposed for all soil layers at each of the sites so that the time span of available data is sufficient to provide reliable results. Even though this leads to variability in the period of reference among the 39 sites, the results are not significantly different when the analysis is performed using a shorter/longer period of data as it will be shown.

In the second part of this chapter (Section 3.2.3), SoTOS is used to explore the relationship between SAT and GST at interannual to decadal timescales. In this part, only the SAT and the ST_{005} variables are employed. The ST_{005} is considered as the GST since this is the closest available level to the surface. First, the structure of the SAT and GST covariance during the whole available period is evaluated. This is achieved by estimating the Pearson's correlation coefficients between SAT and GST. The correlations are calculated using monthly anomalies (i.e., annual cycle removed by subtracting from each monthly value the corresponding monthly long-term average). Also, the differences between the SAT and the GST mean values are evaluated as this gives additional information

about the energy exchange across the air-ground interface. Second, the temporal evolution of the SAT–GST relationship is explored. This is done by estimating the SAT and GST linear trends independently, and the SAT minus GST linear trend. This assessment allows identifying potential SAT–GST decoupling processes on interannual to decadal timescales. In this part of the chapter, the analyses are developed for annual and DJF and JJA seasons. This is because the energy fluxes at the surface can vary significantly across the year.

The trend analysis is performed for different reference periods due to the variability in the data available for both sites and variables among SoTOS. While some of the sites include SAT and GST information over the whole 1985–2018 period, some other sites cover a shorter time frame, particularly for GST (see Figure 2.1). This situation implies that not all of the sites can be employed to analyze the SAT–GST relationship using the full 1985–2018 period. According to this, the sites with available records in SAT and GST covering at least 30 years ending in 2018 are used to evaluate the changes in the SAT–GST relationship over the 1988–2018 period. Sites with at least 20 years of available records are employed to estimate the SAT and GST trends over the 1998–2018 period, and sites with at least 10 years of data are used to evaluate the trends over the 2008–2018 period. The linear trends are calculated in all cases using monthly average anomalies after removing the annual cycle.

The results are compared to the ERA5-L dataset. Therefore, all the analyses described for the SoTOS sites are also implemented in ERA5-L. In order to provide metrics regarding the performance of ERA5-L in representing the SAT and STs of the SoTOS database, Taylor diagrams (Taylor, 2001) are developed. This is a useful way to compare the correlation, standard deviation, and root mean square error (RMSE) between the observation and the reanalysis data. This evaluation is carried out by selecting the nearest grid points in the ERA5-L grid to the observational sites. These grid points are referred to as co-located grid points to the observational sites. The Taylor diagrams are constructed from monthly anomalies (after removing the annual cycle). In this type of diagram, the correlation coefficient is provided on a clockwise angle scale. The radial coordinate gives the standard deviation ratio; high correlation coefficients and standard deviation ratio around 1 indicate low RMSE values (concentric green circles). The comparison of the observations to the reanalysis data are discussed across the chapter.

3.2 Results

3.2.1 SAT and ST climatologies

The use of SoTOS allows providing an overall description of the SAT and ST climatology for the 1985-2018 period over peninsular Spain. Figure 3.1 shows the mean annual temperature for the 39 sites included in SoTOS, as well as the EAR5-land dataset in the background. Within this time, the mean annual temperature is above 10 °C for the bulk of the SoTOS sites. The highest values, of approximately 20 °C, are recorded over the southernmost part of the Iberian Peninsula (e.g., sites 1 and 2) with a gradual decrease towards northern Spain. However, there are some areas over northwest Spain, the Ebro Valley and the Balearic Islands in which the annual mean temperature can be as high as in the southern half of the Iberian Peninsula (> 15 °C; sites 17, 19, 20, and 22). The northern Inner Plateau exhibits, on average, the coldest annual SAT with temperatures around 11 °C. The minimum annual temperature (6.2 °C) is observed at the high elevation of the Sierra de Guadarrama (site-14). ERA5-L shows the same distribution of temperatures with the largest values over the Guadalquivir valley and some areas of the Murcian coast that gradually decrease towards northern Spain. Likewise, the coldest annual mean temperatures are represented over the northern Inner Plateau, the Central System, the Cantabrian Mountains, and the Pyrenees. This spatial pattern of mean annual SAT is also in agreement with the distribution of SAT depicted by some high-resolution temperature observational datasets for the Iberian Peninsula (e.g., Gonzalez-Hidalgo et al., 2015; Serrano-Notivoli et al., 2019).

Interestingly, STs at all depths show overall warmer annual mean temperatures than the SAT at all of the sites. Nonetheless, the same spatial distribution of temperatures is observed, with higher values in the southern half of the Iberian Peninsula, the Ebro Valley, and the Balearic Islands while lower over the northern Inner Plateau with the minimum temperatures in SoTOS at the Sierra de Guadarrama (Figure 3.1). This spatial pattern is shown by ST at all depths. Likewise, the mean annual ST is broadly the same for all of the available soil layers. The distribution of mean annual STs in the ERA5-L agrees with the spatial pattern depicted by the SoTOS sites. However, it is remarkable that the SoTOS sites show slightly higher STs than ERA5-L, especially those located in the northern half of the Peninsula.

Figures 3.2 and 3.3 show the distribution of mean SAT and ST at all depths for the NH winter and summer seasons (DJF and JJA, respectively, hereafter). The DJF shows the coldest values over the northern Inner Plateau that increase both north- and south-wards and over the high mountain areas. An interesting

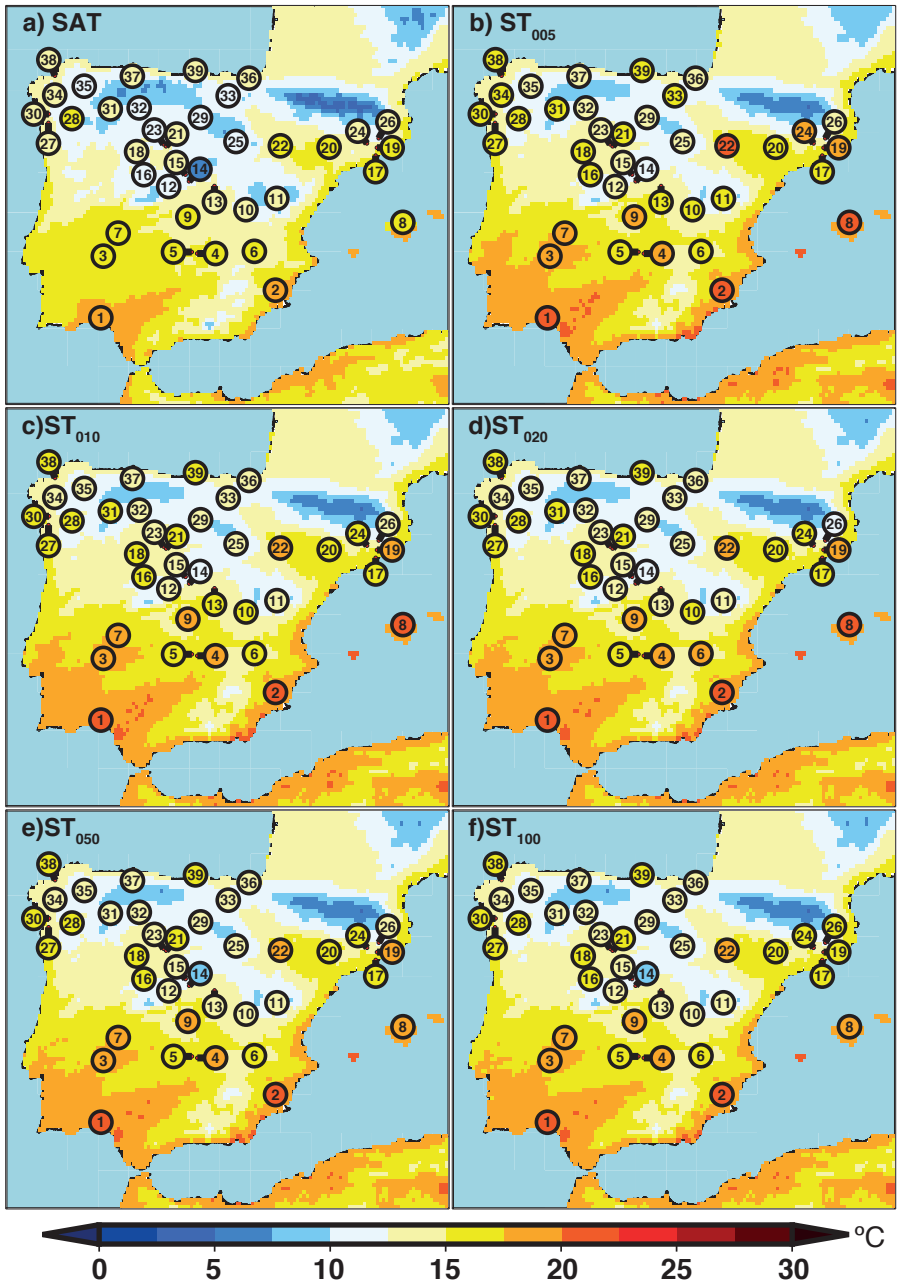

Figure 3.1: Mean annual temperature for SAT, ST_{005}, ST_{010}, ST_{020}, ST_{050}, and ST_{100} (from a to f, respectively). The period employed to obtain the mean values may differ from site to site because of the variability in the available data for both sites and variables. The mean temperature from the ERA5-L reanalysis is shown in the background for the 1985-2018 period. The soil layers of ERA5-L have been interpolated to the depths of SoTOS.

pattern emerges in the DJF data represented by milder winter temperature over the coastal areas compared to the inner continental land, thus reflecting a continentality effect. The ST_{005} shows very similar mean temperatures to SAT with only small differences in some sites (e.g., sites 9 and 22). The mean STs gradually increase with depth and reach values above 10 °C at the deepest layer over the southern half of the Peninsula and the coastal areas. ERA5-L depicts a generally good agreement with the SoTOS sites in the winter months for SAT and ST at all depths.

In JJA (Figure 3.3), there is a more evident southwest-northeast gradient with the highest temperature values over the southern half of the Iberian Peninsula that decreases towards northern Spain. As in the annual average case, the ST at the upper soil layers tends to be warmer than SAT. Note that the ST_{005} shows much larger values than the SAT at all locations. The nature of these differences are discussed in Section 3.2.3. ERA5-L also shows higher mean STs than SAT in the summer months; however, the mean ST at all depths is generally milder compared to the SoTOS sites.

Figure 3.4 shows Taylor diagrams for SAT, ST_{005}, and ST_{100} as an example. High similarity between SoTOS and ERA5-L can be observed for SAT in the annual case as well as in JJA and DJF, with the bulk of the sites showing statistically significant ($p< 0.05$) correlation coefficients above 0.9 (Figure 3.4a and d) and with the standard deviation ratio being distributed around 1. Therefore, there is a good capability of ERA5-L to reproduce the SAT observations. For the ST_{005}, there is a lower skill in reproducing the observations as the correlation coefficients are lower compared to SAT for the annual case, although most of them are still high (>0.7; $p< 0.05$). Similarly, the standard deviations are systematically below 1 revealing that EAR5-L underestimate the variability of observations. The JJA and DJF cases suggest that most of the annual average variability for ST_{005} is influenced by the JJA season as the data in this season are spread towards lower correlation coefficients and standard deviation ratios are below 1, whereas in DJF, correlation is higher and standard deviation ratios cluster around 1 in most of the cases. A similar picture to the ST_{005} is depicted for the ST_{100}. The correlation coefficients range from 0.4 to 0.9 in the annual data, and the standard deviation ratios are distributed below 1. The JJA and DJF seasons show a wider range of correlation values ranging from 0.2 to 0.95 and standard deviation ratios below the 0.5 arc for the JJA season. The apparent reduction in the ability of the ERA5-L to reproduce the soil thermal state over the SoTOS sites, with depth may be related to different issues, including the spatial resolution of the model and the interpolation process (Yang and Zhang, 2018). Nevertheless, it may also be related to assimilating only surface data (Balsamo et al., 2015) or to the use of a shallow BBCP in the model (Smerdon and

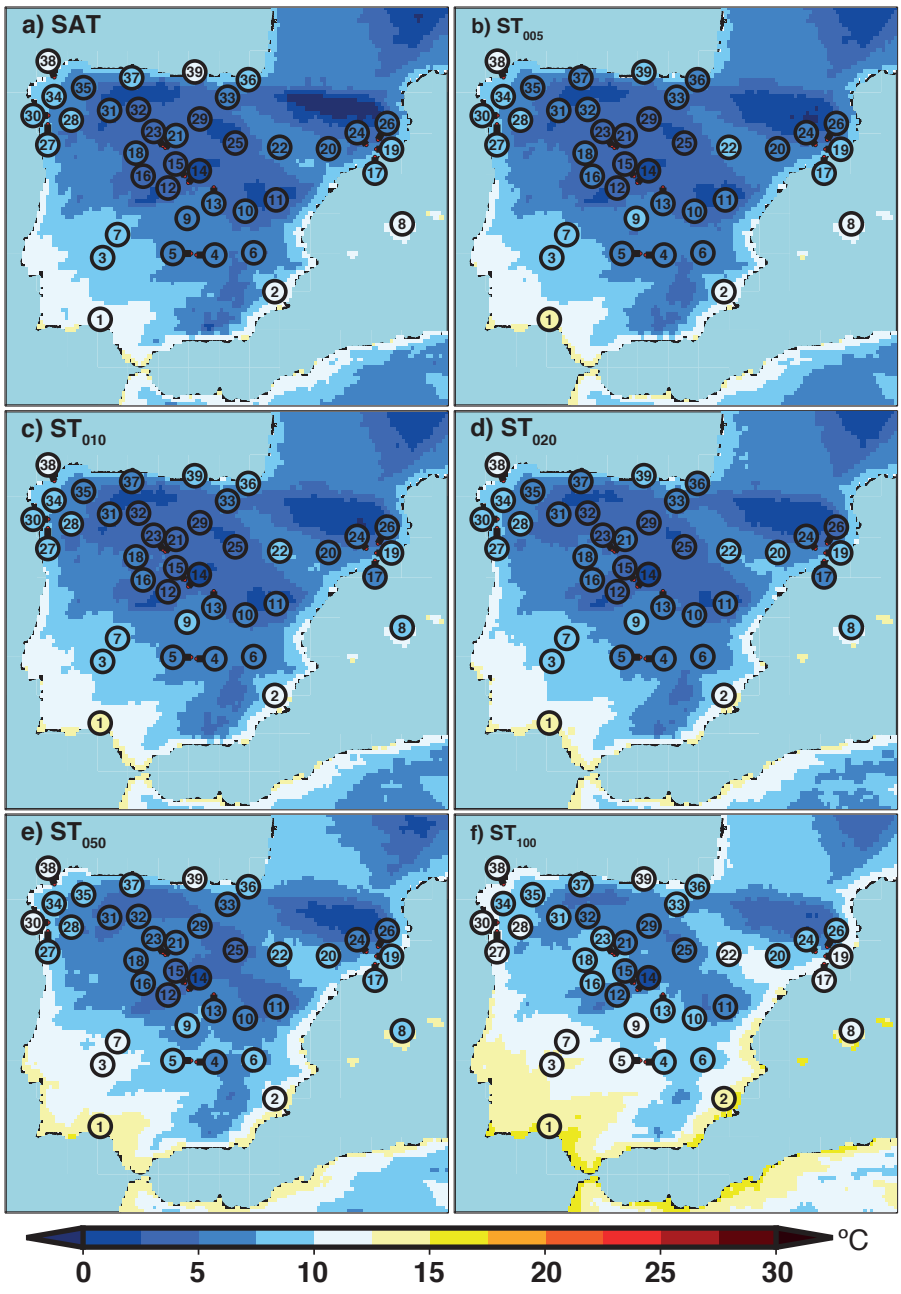

Figure 3.2: As Figure 3.1 but for DJF.

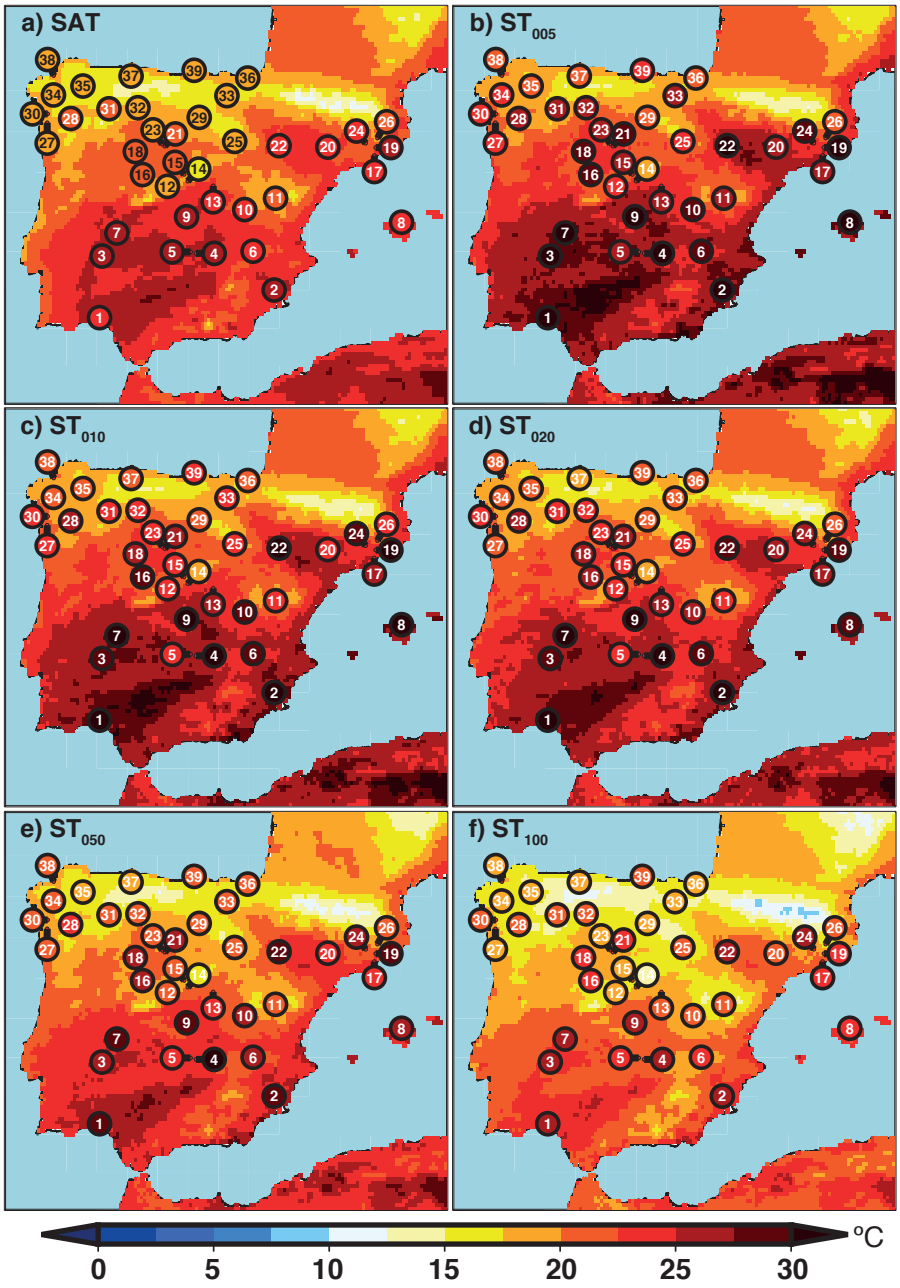

Figure 3.3: As Figure 3.1 but for JJA.

Stieglitz, 2006). The worse values in JJA may be related to a variety of reasons like radiation-vegetation interactions or irrigation (Georgescu et al., 2008). Up to this point, exploring it further is beyond the purpose of this work.

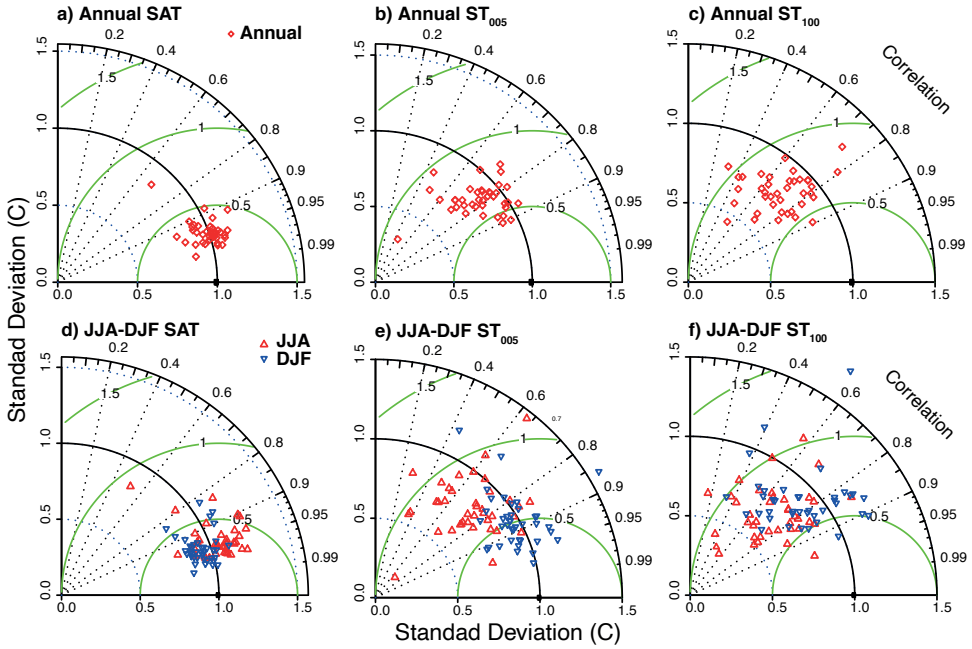

Figure 3.4: Taylor diagrams of the monthly anomalies at each site for the annual SAT (a), ST_{005} (b), ST_{100} (c) as well as JJA and DJF for SAT (d), ST_{005} (e) and ST_{100} (f).

3.2.2 Subsurface thermal regime

Figure 3.5a illustrates the time series of daily SAT and ST data for one of the SoTOS sites (site-2). Note that there is no continuous data coverage for the whole 1985–2018 period for ST at all of the layers. Despite such discontinuities, there are some periods in which data is available at all depths (e.g., 1997–2015 for site-

2). These periods of corresponding coverage of ST data permit the extraction of both the amplitude and phase of the different soil layers in a homogeneous frame, so the analysis of the conductive heat transport of the annual temperature signal into the subsurface is not biased for gaps in the data. All the SoTOS sites contain ST data at all depths for corresponding periods of at least five consecutive years.

The amplitude and phase of the annual signal for SAT and ST at all depths are extracted using a linear least-squares fitting, as described in Section 3.1. Light-green lines in Figure 3.5a depict the regression results for each variable for the periods in which the original data is available. The amplitude and phase information is employed to represent the annual cycle of temperatures (Figure 3.5b). Interestingly, the SAT annual signal shows a lower amplitude relative to the ST at the upper layers. This feature is common to the majority of the sites analyzed in the present study. The nature of such a response is discussed in Section 3.2.3. A simple visual inspection reveals the amplitude attenuation and phase shift with depth of the downward propagating wave, visible, for example, in ST_{100} relative to the upper layers. Thus, qualitatively supporting the characteristic of the conductive heat transport of the surface temperature signal into the soil. However, a more quantitative analysis is required to provide a reliable characterization of the shallow soil thermal regime.

Accordingly, the variation with depth of the natural logarithm of the amplitude (Ln-A, hereafter) and the phase shift (P-S, hereafter) for the available ST layers is considered and a linear regression model is applied to evaluate the relationship between either the Ln-A or P-S shift with depth. Figure 3.6 shows the results of this analysis for sites 2, 14, 21 and 30 as an example. These sites have been selected because they are representative of different areas of peninsular Spain, such as the Mediterranean coast, the high elevation of the Central System, the northern Inner Plateau, and the northwestern coast over the Atlantic ocean, thus including different climatic and soil conditions. The regression lines in Figure 3.6 clearly show the linear dependence of both Ln-A and P-S with depth. In all cases, the r^2 of either the Ln-A or P-S is above 0.95. As indicated in Section 3.1, the slope represents the wave vector of the propagating thermal wave. This information is used to estimate the apparent thermal diffusivity and obtain a metric of conductive heat transport. The estimated thermal diffusivity from both the Ln-A and P-S is indicated in each of the panels.

Intriguingly, the estimations from the Ln-A and P-S linear regressions can often be different. Indeed, this is the general behavior in the bulk of the sites included in SoTOS. In some cases, such difference is relatively small within the range of a few tenths of $\times 10^{-7} m^2 s^{-1}$ (e.g., sites 2 and 30) while some other cases show larger differences (e.g., sites 14 and 21). In previous works that have used a similar approach to estimate the mean subsurface thermal diffusivity, the

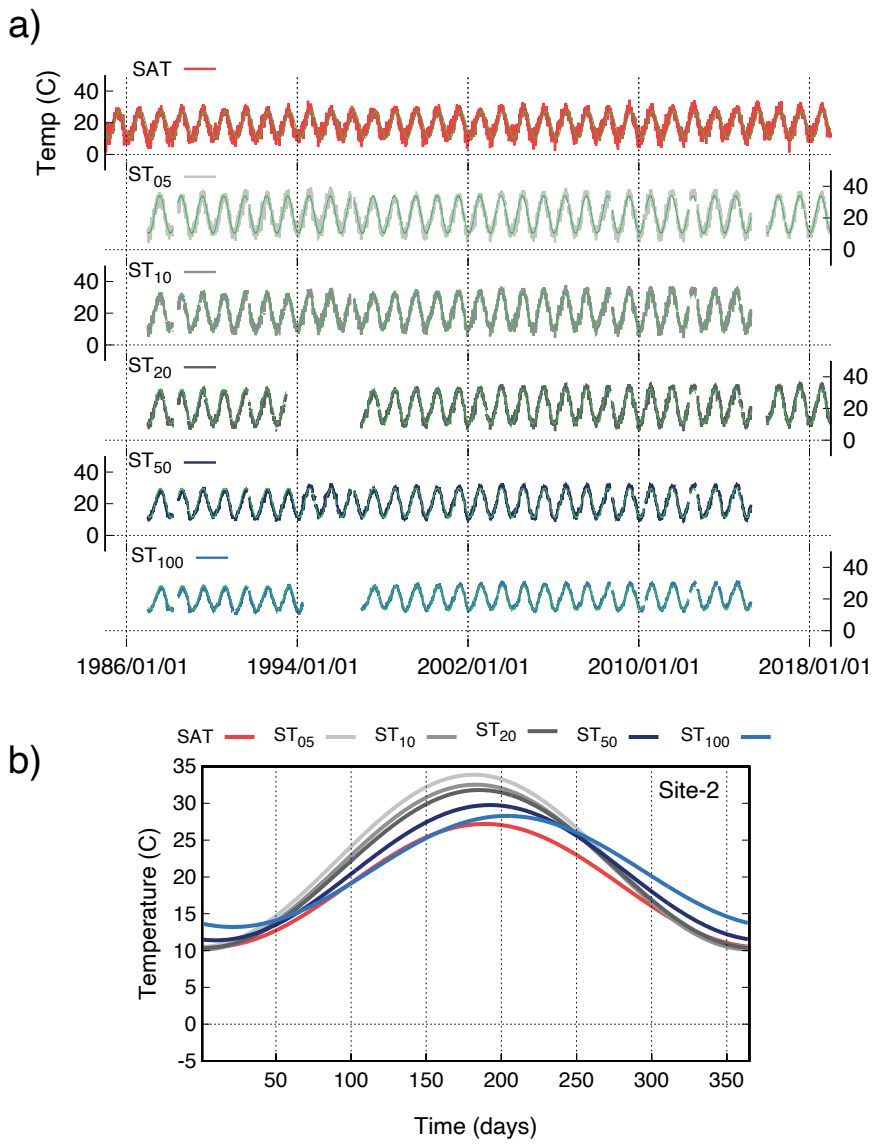

Figure 3.5: Final quality controlled daily SAT and ST_{005}, ST_{010}, ST_{020}, ST_{050} and ST_{100} for site-2 as an example. The light-green line in a) indicates the results of fitting the data to a sine function. Additionally, the annual cycle for SAT and ST at all depths is shown in b).

values obtained from either the Ln-A or the P-S closely coincide (e.g., Smerdon et al., 2003, 2004). However, SoTOS differs from those case studies by including ST observations only for the subsurface upper meter, whereas in other works, the soil temperature profiles go up to at least 3 m depth and reach >10 m. In such cases, the bottom part of the soil profiles is less affected by soil-hydrology-related processes such as evapotranspiration or soil freezing and thawing that may have a considerable influence on altering the pure conductive regime or disrupting it.

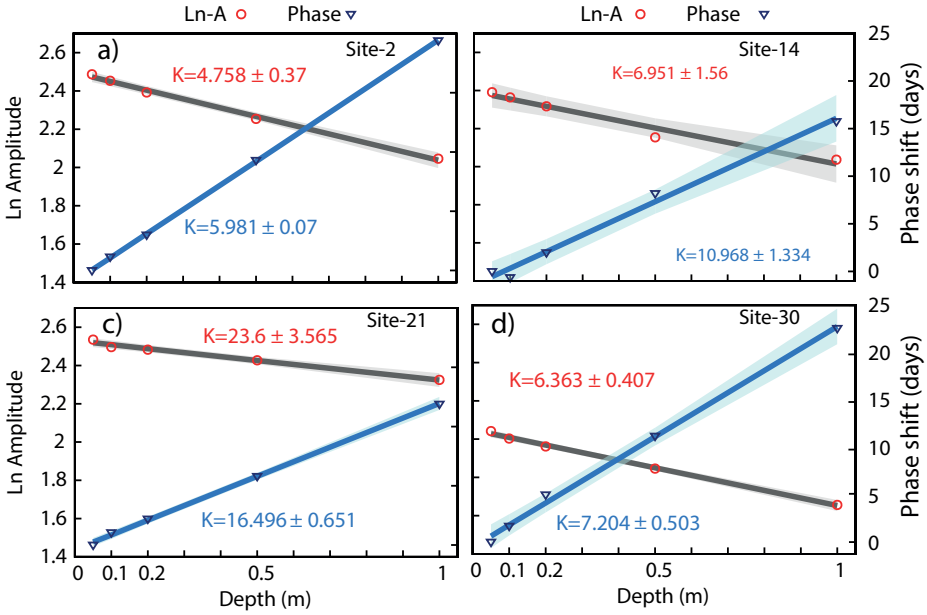

Figure 3.6: Linear regression of the natural logarithm of the amplitude (red circles) and phase shift of the annual signal of STs at the levels with available data (blue triangles) relative to the ST_{005} as a function of depth for sites 2, 14, 21, and 30 as an example. The calculated apparent thermal diffusivity from both the natural logarithm of the amplitude (grey) and phase shift (blue) are shown in each panel. The units for thermal diffusivity are given as $\times 10^{-7} m^2 s^{-1}$. The regression fit (line) and associated uncertainties are also shown.

Indeed, we assume a vertically homogenous soil with constant thermal properties through time in our estimation of the apparent thermal diffusivity. Further, we neglect soil-hydrology-related processes as we take into account only thermal

conduction (Equation 3.2). Gao et al. (2008b) analyzed the effects of liquid water flux into the soil on estimating the apparent thermal diffusivity at a daily temporal scale in a site over the Loess Plateau of China. They showed that estimating the shallow soil apparent thermal diffusivity from the amplitude attenuation or the phase shift with depth of the daily cycle, considering only thermal conduction, yield slightly different results. However, if the effect of convection is taken into account, an improved estimation of the apparent thermal diffusivity can be achieved. Tong et al. (2017) followed up on Gao et al. (2008b) approach and showed that the estimation of the amplitude and phase of the daily cycle at different depths is highly sensitive to small changes in liquid water flux. Thus, influencing the estimation of the apparent thermal diffusivity when only conduction heat transfer is considered. Such findings highlight the contribution of soil-hydrology-related processes on the thermal regime within the shallow subsurface. Although their results are based on the daily cycle, the overall effect should be similar on an annual scale. In both cases, the amplitude and phase shift of soil temperature at different depths are extracted assuming that soil temperature varies sinusoidally.

Figure 3.7 shows the spatial distribution of the apparent thermal diffusivities obtained from both the Ln-A (a) and P-S (b) vs. depth for both SoTOS and ERA5-L. For the ERA5-L, the estimation of the apparent thermal diffusivity has been performed using the interpolated layers to the SoTOS depths, so in both cases, the thermal diffusivity represents the mean value over the upper 1 m depth. In most of the SoTOS sites, the apparent thermal diffusivity obtained from the Ln-A vs. depth is lower than the one obtained from the P-S, although there are also a few cases in which larger values arise from the phase shift (reddish circles). Despite this, the sites indicating higher values are generally the same in both cases. The spatial variability shows larger values over the northern coastal edge of the Bay of Biscay (e.g., sites 36, 37 and 39) and some other sites dispersed over the Peninsula (e.g., sites 3, 21 and 25). Interestingly, a similar spatial pattern is observed in the ERA5-L data, particularly in the case of the Ln-A (Figure 3.7a), with the largest values along the northern coast and the border with France. This spatial pattern is also evident in the case of the P-S for the EAR5-L data. However, there is lower spatial variability. This is because the amplitude and phase of the annual signal are obtained from data at monthly resolution. Since the phase shift at these depths is on the order of days, it cannot be resolved precisely. In the case of the ERA5-L data, the estimation of the apparent thermal diffusivity yields lower values than those obtained from the observations. Therefore, the results have been scaled by a factor of 2 to improve spatial patterns visualization.

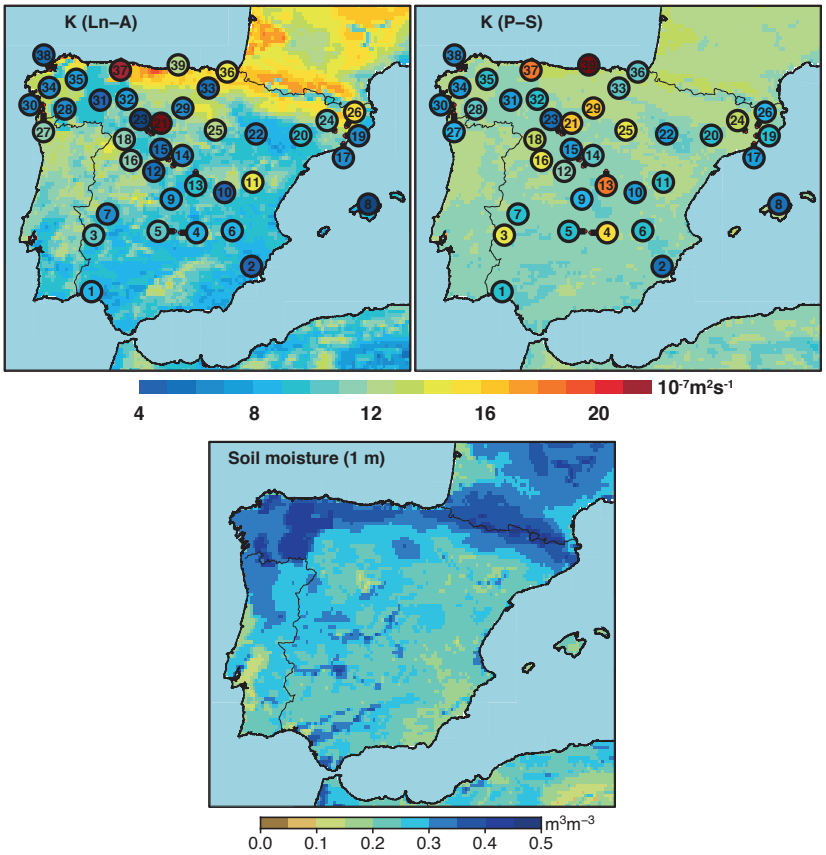

Figure 3.7: Soil apparent thermal diffusivity calculated from the slope of the linear regression of the natural logarithm of the amplitude (a) and phase shift (b) as a function of depth for the 39 SoTOS sites. The results from ERA5-L are shown in the background and have been scaled with a factor of 2 for visibility. (c) ERA5-L average volumetric water content for the upper subsurface meter was obtained as a weighted average over the layer thickness.

The spatial variability in the estimated apparent thermal diffusivity should be related to the soil organic and mineral composition (Farouki, 1986) and the soil water content because it affects the soil thermal conductivity (Patil et al., 2011) and the volumetric heat capacity (McCumber and Pielke, 1981). Figure

3.7c shows the ERA5-L volumetric water content of the upper subsurface meter as a weighted average over the layer thickness. The spatial distribution reveals a direct relationship between soil water content and soil thermal diffusivity in the reanalysis data. Note that larger thermal diffusivities coincide with larger soil water content. This is especially evident in the Ln-A map, but such a signal is also detected in the P-S map. Although other factors also influence soil thermal diffusivity, the soil water content explains the spatial distribution observed to a large extent.

Comparing the evolution of the apparent thermal diffusivity to the variations in soil moisture in a time series analysis can provide additional information regarding the influence of soil water content on the thermal properties of the soil. Although there is no available direct information on the soil water content for the sites included in SoTOS, the ERA5-L data can provide insights into the variations with time of thermal diffusivity and its relationship to soil water content changes. To estimate the variations in the apparent thermal diffusivity with time, the SoTOS sites with the longest consecutive ST records at all depths are selected. This results in a 20-yr time span, from 1994 to 2014, for sites 7, 11, 16, 21, 23 and 32. The apparent thermal diffusivity is estimated for 18 chunks of 3 consecutive years within the 1994–2014 period using a 1-yr moving window employing the same approach described in Figure 3.6. Therefore, we obtain 18 consecutive estimates of the thermal diffusivity. The same approach is applied to ERA5-L using the co-located grid points to the observational sites in SoTOS over the 1981–2018 period. In this case, the entire soil model profile is employed, and the reference depth for the soil temperature to evaluate the amplitude attenuation and phase shift with depth is considered as the midpoint of each soil layer (i.e., 0.035, 0.175, 0.64 and 1.945 m). The co-located grid points volumetric water content is used to assess the influence of soil moisture on the variability in thermal diffusivity with time. The volumetric water content is averaged over three consecutive years to match the estimates of the thermal diffusivity periods.

Figure 3.8 illustrates the evolution of the mean apparent thermal diffusivity from SoTOS and ERA5-L and the ERA5-L soil moisture for sites 11 and 21 as an example. Only the Ln-A results are depicted because of the above-mentioned limitations to obtain the phase from monthly data in the case of ERA5-L. The apparent thermal diffusivity from ERA5-L is scaled by a factor of 2 to improve the visualization of the results. The soil moisture is shown as the weighted average over the upper meter of the soil (i.e., L1, L2 and L3) to compare with the results from SoTOS. Also, a weighted average over the entire soil profile (i.e., L1, L2, L3 and L4) is shown to assess the thermal diffusivity-soil moisture relationship from solely model data. The apparent thermal diffusivity shows interannual variability in both SoTOS and ERA5-L with larger diffusivity changes in So-

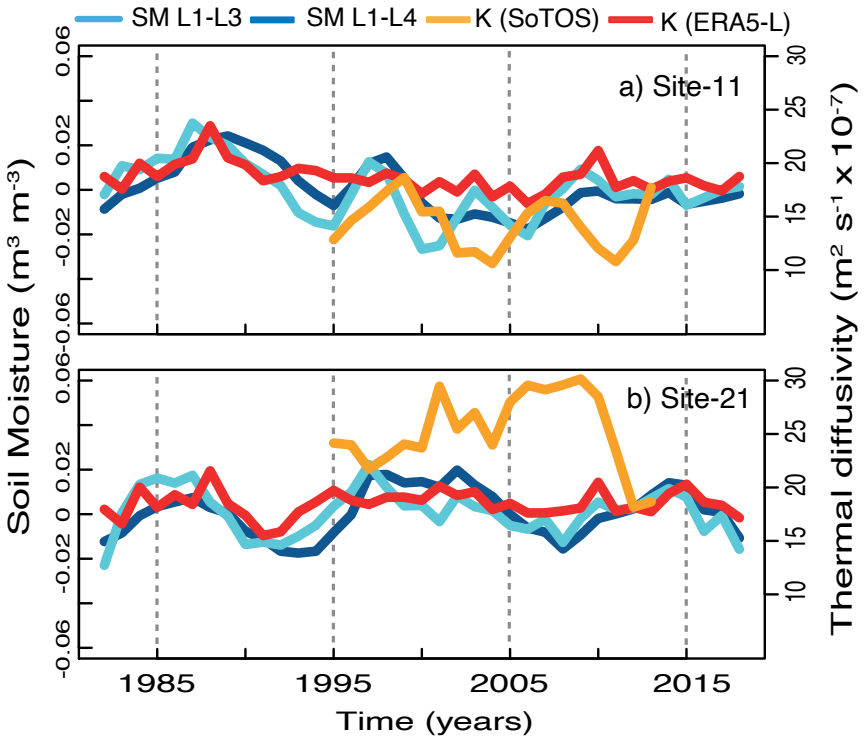

Figure 3.8: Evolution of apparent thermal diffusivity in SoTOS (ERA5-L) estimated for 18 (35) different chunks of three consecutive years within the 1994–2014 (1981–2018) period using a 1-yr moving window. The apparent thermal diffusivity is presented only for the Ln-A analysis as described in Figure 3.6. The entire soil model profile is employed to estimate the thermal diffusivity of ERA5-L. The reference depths to evaluate the relationship between amplitude attenuation and phase with depth are the mid point of each model soil layer (i.e., 0.035, 0.175, 0.64, and 1.945 m depth). The apparent thermal diffusivity from ERA5-L is scaled by a factor of 2 for visibility. The ERA5-L volumetric water content is also shown for the 1981–2018 period as a weighted average over the upper meter of the soil (i.e., L1, L2 and L3) and also as a weighted average over the entire soil profile (i.e., L1, L2, L3 and L4). Results are shown for sites 11 and 21 as an example.

TOS. In the SoTOS data, there is no clear relationship with the ERA5-L soil moisture. Nonetheless, this can be an expected result due, in part, to the model spatial resolution and model parameterizations or bias. On the contrary, when this analysis is developed using solely model output, a clear relationship between thermal diffusivity variability with time and soil moisture content emerges. In both examples, there is a positive and significant correlation (0.67 and 0.53 for sites 11 and 21, respectively; $p < 0.05$) between the apparent thermal diffusivity and soil moisture content. Accordingly, the results show a consistent response of thermal diffusivity to changes in soil water content.

Despite the discussion above regarding the differences between the Ln-A and P-S apparent thermal diffusivity estimates, the resulting values, in either case, agree with *in situ* measurements in different parts of the world. For instance, Oladunjoye and Sanuade (2012) measured the thermal properties of soil, including the thermal diffusivity, in the area of Ogun state (Southwestern Nigeria) for the upper 1.5 m soil depth. Their measured thermal diffusivity range from 3.46 to 7.5 $\times 10^{-7} m^2 s^{-1}$ for sandy soils with relatively low soil moisture content. Similarly, Sáez Blázquez et al. (2017) developed a thermal conductivity (α) map for the Avila region (Spain) by measuring different soil thermal conductivities, including sandy, clayey and mud-like soils. They found that soil thermal conductivity over this area ranges from ∼1.44 to 2.5 W/mK depending on soil type and moisture conditions. Assuming typical values of density (ρ) and specific heat (c_p) for soil mineral and soil organic matter around 1800 Kg/m^3 and 1400 J/KgK (Kluitenberg, 2002), these conductivity values would roughly correspond to thermal diffusivities of 6–10 $\times 10^{-7} m^2 s^{-1}$ ($\kappa = \alpha / c_p \rho$). This range of variations is similar to the apparent thermal diffusivity obtained from the analysis developed in the present work (Figure 3.7).

Even though non-conductive processes in the shallow soil may alter the pure conductive heat transport over the SoTOS sites, the conductive regime does explain the largest part of the temperature distribution in the upper meter of the subsurface, as it has been shown. To further illustrate this issue, the steady-state analytical solution of the one-dimensional heat conduction equation for a simple harmonic signal (Equation 3.2) is used to simulate the annual cycle of soil temperature for the SoTOS soil layers at the 39 sites. This is achieved by using the estimated amplitude and phase of the ST_{005} as the boundary conditions. The thermal diffusivity is that resulting from the linear regression of both the Ln-A and P-S vs. depth for comparison. The results are then compared to the observed annual cycle.

Figure 3.9 depicts the observed and estimated annual cycle (solid and dashed lines, respectively) for the ST_{005}, ST_{010}, ST_{020}, ST_{050} and ST_{100} for site-21 as an example. Figure 3.9a shows the results when the apparent thermal diffusivity

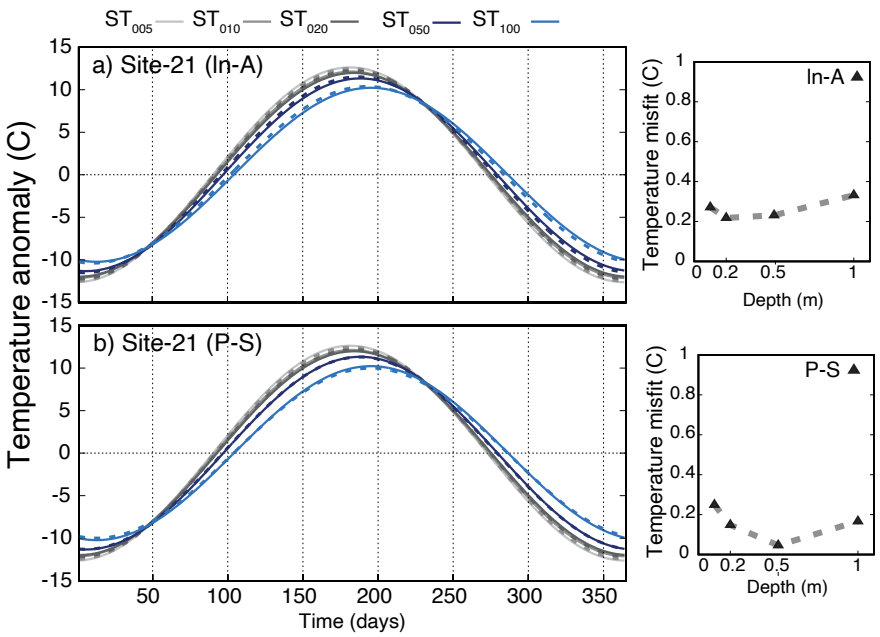

Figure 3.9: Observed annual cycle for ST at all of the layers (solid line) and estimated annual cycle (dashed lines) by propagating the ST_{005} into the subsurface using the steady-state analytical solution of the one-dimensional heat conduction equation for a simple-harmonic surface temperature signal (Equation 3.2). The synthetic annual cycle is estimated using the apparent thermal diffusivity obtained from the Ln-A (a) and P-S (b). The right-hand side panels show the RMSE between the observed and simulated data. The results are shown for site-21 as an example.

estimated from the Ln-A approach is employed, whereas Figure 3.9b uses the apparent thermal diffusivity estimated from P-S. Note that in both cases, the observed and estimated data are nearly identical. Although the differences in the calculated thermal diffusivity can be significant within the range of uncertainties at some locations, the propagation of the ST_{005} signal to deeper layers yields very close estimations to the observed annual cycle even in the cases with the largest differences. The RMSE between the observed and simulated annual cycle (right panels) shows that the temperature misfit is slightly larger for P-S than Ln-A but relatively low in either case. These results support the premise of conductive

heat transport controlling the propagation of the surface temperature signal into the subsurface. Further understanding of the non-conductive processes involved at the different locations of our ST dataset that produce the differences described in this section would be desirable. However, this requires a more detailed description of site-specific soil types, vegetation cover and soil hydrology characteristics, among others. Such a level of information is not available at the very local scale that is required.

3.2.3 SAT-GST coupling

The SAT–GST coupling is usually assessed by analyzing the differences between SAT and GST mean values as well as their covariance structure over a reference period of time (Bartlett et al., 2004; González-Rouco et al., 2009; Melo-Aguilar et al., 2018). Both of these analyses provide information on the interactions across the air–ground interface. Figure 3.10 shows the SAT-GST differences (left) and the correlation coefficients (right) for the sites included in SoTOS. The results are shown for DJF (top) and JJA (middle) seasons separately in addition to the annual average (bottom). The SAT-GST differences are calculated from absolute temperature data using the original daily values. For the estimation of the correlation coefficients, monthly averages are employed from which the annual cycle has been removed by subtracting the corresponding monthly long-term average from each monthly value over the whole available period. A few cases exist in which the existence of missing data in GST either in DJF or JJA over relatively large periods prevents from calculating the correlation coefficients (e.g., sites 6, 14, 26, 38 and 29). Therefore, these sites are excluded from the analysis. The ERA5-L dataset is also employed to compare the results of the SoTOS sites along this section.

Figure 3.10 shows smaller differences in DJF compared to JJA. In DJF, low non-significant positive and negative values are observed for most of the SoTOS sites. Interestingly, the bulk of sites located over coastal areas show positive SAT-GST differences, and in some cases, relatively large and significant values are observed (sites 17, 36 and 37). This effect of warmer SAT relative to GST over some coastal areas in the winter months has also been reported over North America's eastern coast (Smerdon et al., 2004). Similarly, Melo-Aguilar et al. (2018) showed with the use of ESM output that the influence of the ocean mainly drives such a response as some energy is released from the ocean surface to the air above due to the temperature difference in the winter months, thus warming the air with a net effect of higher SAT relative to GST. The DJF correlation map

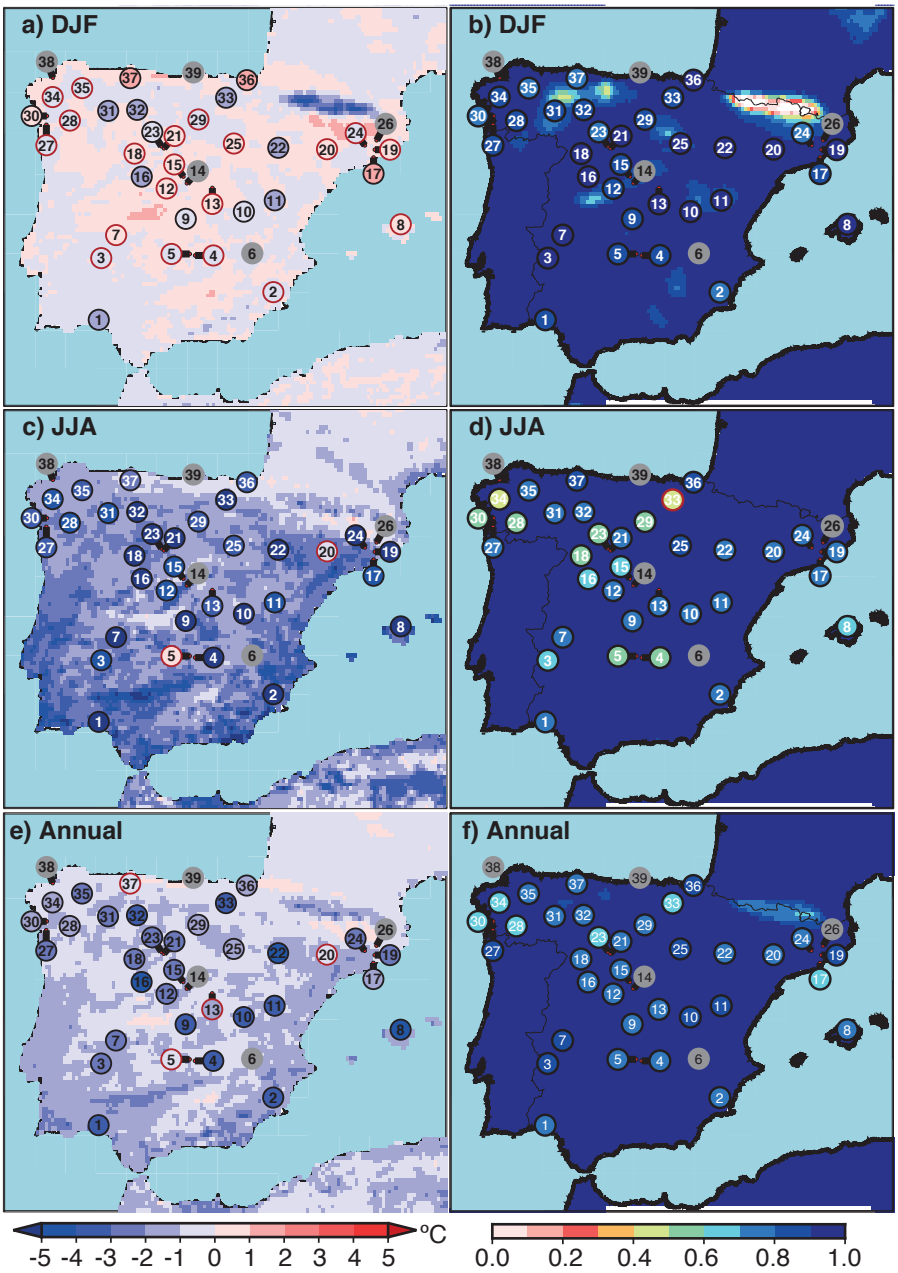

Figure 3.10: (a) DJF, (c) JJA and (e) annual SAT-GST differences at the 39 sites. (b) DJF, (d) JJA and (f) Annual correlation coefficients between SAT and GST at the 39 sites. The period of reference to estimate the differences and correlations may differ from site to site, depending on the available data for each of the sites. Black (red) contour circles indicate statistical significant (non-significant) values ($p < 0.05$). The grey-colored sites represent those locations with no sufficient data to estimate the correlations. The results from the ERA5-L dataset are shown in the background.

shows values above 0.9, thus denoting a strong SAT–GST coupling in this season. The ERA5-L data agree in representing the SAT-GST differences and correlations in the DJF months. Remarkably, the reanalysis correlations show a general distribution of values above 0.95 except in the mountain areas. Particularly, the low correlation coefficients over the Pyrenees are worth noting. A similar pattern with high negative SAT-GST differences is observed over this region, which is likely the response to the snow cover insulating effect that prevents the contact of the cold air with the ground surface (Melo-Aguilar et al., 2018). SoTOS does not contain information over this region or other high elevation areas to evaluate such an effect in observational data. Exploring this effect using SAT and GST observations can provide further insights into these mechanisms.

The JJA maps show different behavior in comparison to DJF since the majority of SoTOS sites show negative SAT-GST differences with some large values distributed over both the northern and the southern plateau, as well as over coastal areas of the Mediterranean sea and the Atlantic ocean that can reach $>$ 7 °C in absolute value. Likewise, the correlations coefficients decrease for some of the SoTOS sites, although they are still high (>0.5) in all cases. ERA5-L shows a good agreement with the SAT-GST differences in the summer months with negative differences that coincide with those of SoTOS. However, the SAT-GST differences are of smaller magnitude compared to SoTOS. Other works that have also compared the ERA5-L reanalysis and its previous version (ERA-Interim/land; Balsamo et al., 2015) to soil temperature observations over China also indicate that these reanalysis products tend to underestimate soil temperatures, especially during summer (Yang and Zhang, 2018; Li et al., 2020). Likewise, ERA5-L does not represent the lower correlation depicted by some SoTOS sites in this season since high correlations are widespread in the reanalysis. The cases in which SoTOS and ERA5-L do not agree in the correlation patterns during JJA may be related to local processes due to site-specific characteristics that the ERA5-L reanalysis cannot detect because of the spatial resolution.

The annual average indicate a combination of both seasons. It shows a predominant distribution of warmer GST relative to SAT for both SoTOS and ERA5-L with negative differences that range from 0 to -5 °C. The largest differences appear to be distributed over the Inner Plateau, although there is no clear spatial pattern since relatively large values can also be found over the Mediterranean coats, the Ebro Valley and the Balearic Islands, as in JJA. The map of correlations indicates a general distribution of high and significant values ($p< 0.05$). Interestingly, the lower correlations over the Pyrenees depicted in DJF in the ERA5-L data are also observed in the annual average. Thus, highlighting the influence of the winter season on the annual average over this area.

The combined analysis of SAT-GST differences and correlations indicates a strong SAT-GST coupling over Spain. However, the smaller (higher) SAT-GST differences (correlations) in DJF relative to JJA suggest a stronger bond in DJF compared to JJA. Larger SAT–GST offset in JJA relative to DJF has been reported by Putnam and Chapman (1996) over an arid region of northwest Utah. They showed that radiative heating of the ground surface during the summer months leads to large SAT–GST offset from March through October. They found that the annual cycle of GST minus SAT differences closely tracks the annual solar cycle. As the solar radiation decreases in the winter months, the differences between SAT and GST average to zero. A similar physical response may explain the SAT–GST relationship over some sites included in SoTOS, especially those located over the Inner Plateau, the Ebro Valley and the coastal areas of the Mediterranean sea and the Atlantic ocean. The latter is due to the high solar irradiance in the summer months (Sancho-Avila et al., 2012) and the low precipitation rates over the mentioned areas (Serrano-Notivoli et al., 2017). These characteristics are like those described for the analyzed observatory in Putnam and Chapman (1996).

The land surface characteristics and the soil moisture content are additional factors contributing to the SAT–GST offset as they control the radiative heating of the land surface: a dry, bare surface would warm more than a vegetated covered area; thus, having a much larger SAT–GST offset (Cermak and Bodri, 2018). The location of most of the SoTOS sites corresponds to bare soil, short grass and shrub surfaces, as shown in Figure 3.11a, consistent with larger SAT–GST offset during the months with the higher incoming solar radiation. The soil moisture content from the ERA5-L provides additional support to the described mechanism. Figure 3.11b shows the mean JJA volumetric water content over the 1985–2018 period for the first model layer. Note that the areas with low (high) water content in this layer closely coincided with the areas of high (low) SAT-GST differences in the JJA map (Figure 3.10c). Indeed, the JJA SAT–GST differences and volumetric water content for the first model layer have a spatial pattern correlation of 0.67.

A comparison between the annual cycle of SAT and GST further illustrates the larger contribution of JJA to the annual average SAT–GST offset. Figure 3.12 shows the observed annual cycle for both SAT and GST in SoTOS using the amplitude and phase of the annual signal obtained from the least-squares fit of the observations to a sine function (see Section 3.1). Likewise, the SAT and GST differences over the year are presented (the shaded region at the bottom panels). The results are shown for sites 2 and 30 as an example. Note that the SAT-GST differences begin to grow in both cases, in absolute values, from the first quarter of the year and reach the maximum difference approximately in the middle. A more pronounced effect is observed in site-2 than in site-30, presumably due to

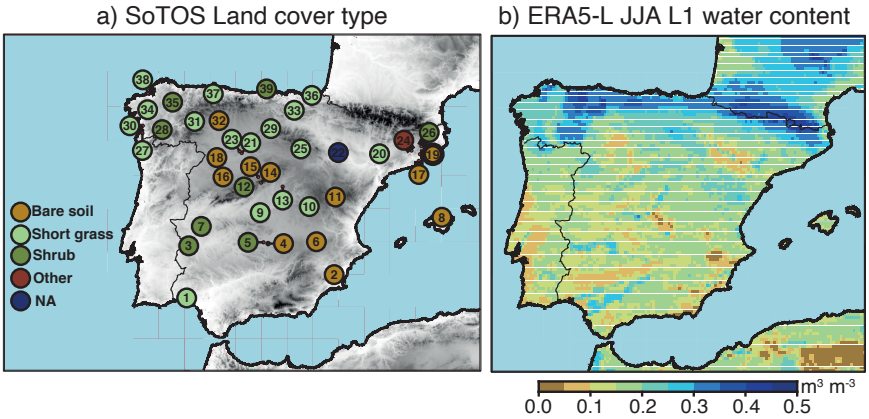

Figure 3.11: Land cover types for the SoTOS locations (a) and mean volumetric water content for the first ERA5-L soil layer over the 1985–2018 period (b).

the dryer conditions and higher solar irradiance of site-2 compared to site-30. In DJF, site-2 shows an SAT-GST offset that averages zero, whereas, in the case of site-30, small positive differences can be observed.

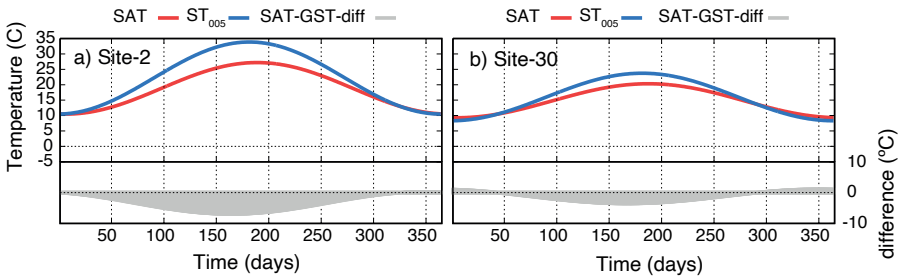

Figure 3.12: Observed annual cycle for SAT and GST (ST_{005}), as shown in Figure 3.5. The grey-shaded area at the bottom represents the SAT-GST differences. Sites 2 (a) and 30 (b) are presented as an example.

At this point, the question of seasonal effects on the SAT–GST relationship at longer timescales is remaining. If the seasonal SAT–GST offset does not change from year to year, the effect would be a constant bias with no consequences on

the SAT–GST tracking at longer than annual timescales (Bartlett et al., 2005). However, if the relative influence of seasonal processes changes over long periods, the long-term SAT–GST coupling may change (Smerdon et al., 2006). In the present study, SoTOS allows for evaluating the SAT–GST relationship over a three-decade period for the sites with the most extensive length in observations for both SAT and GST. In some other cases with shorter temporal records of ST data, it is possible to employ either 20-yr and 10-yr periods. Different reference periods were used: 1988-, 1998- and 2008-2018, respectively. The SAT and GST linear trends for the annual data and DJF and JJA seasons are estimated to assess their evolution at interannual to multi-decadal timescales. Figure 3.13 shows the distribution of trends for both SAT (left panels) and GST (right panels) for the three different periods of reference (1988-, 1998- and 2008-2018 from top to bottom, respectively). The results from ERA5-L are shown in the background. Note that the number of sites depicted in each of the panels increases as the reference period is shorter since more sites include ST data closer to the present. The grey-colored sites represent those locations with no sufficient data to estimate trends. Although for SAT, all of the sites included in SoTOS have complete data coverage for the whole 1985-2018, the trends are presented only for the sites in which there is also available observations for GST to allow for comparison. Overall, positive trends are observed in the SAT for each reference period in both SoTOS and ERA5-L, showing statistically significant values in some SoTOS sites ($p < 0.05$). To evaluate the significance of trends, a t-test has been applied; temporal autocorrelation has been considered for the estimation of both the standard deviations of the regression residuals and the reduced degrees of freedom using a lag-1 autoregressive statistical model (Santer et al., 2000; Hartmann et al., 2013). The trends are systematically higher in the last decade than the 20-yr and 30-yr periods with the largest SAT increases over the northern Inner Plateau and the Mediterranean coast with values that can be >1 °C decade^{-1} (e.g., sites 2, 17, and 21). The ERA5-L reanalysis accurately represents the pattern of SAT trends in SoTOS.

The GST trend maps also show a predominant picture of positive trends, especially in the last 2 decades. However, there are some differences relative to SAT for some of the SoTOS sites, particularly in the 30-yr period. For instance, sites 29 and 30 show negative and significant trends. Likewise, a pattern of larger GST increase relative to SAT emerges in all of the reference periods. The ERA5-L data also show widespread GST warming over Spain. However, although it agrees in identifying the maximum areas of warming, it underestimates the GST warming shown by some of the SoTOS sites (5, 21, 23 and 28).

The trend analysis for DJF (Figure 3.14) and JJA (Figure 3.15) provides insights into the mechanisms that control annual data response. While DJF shows

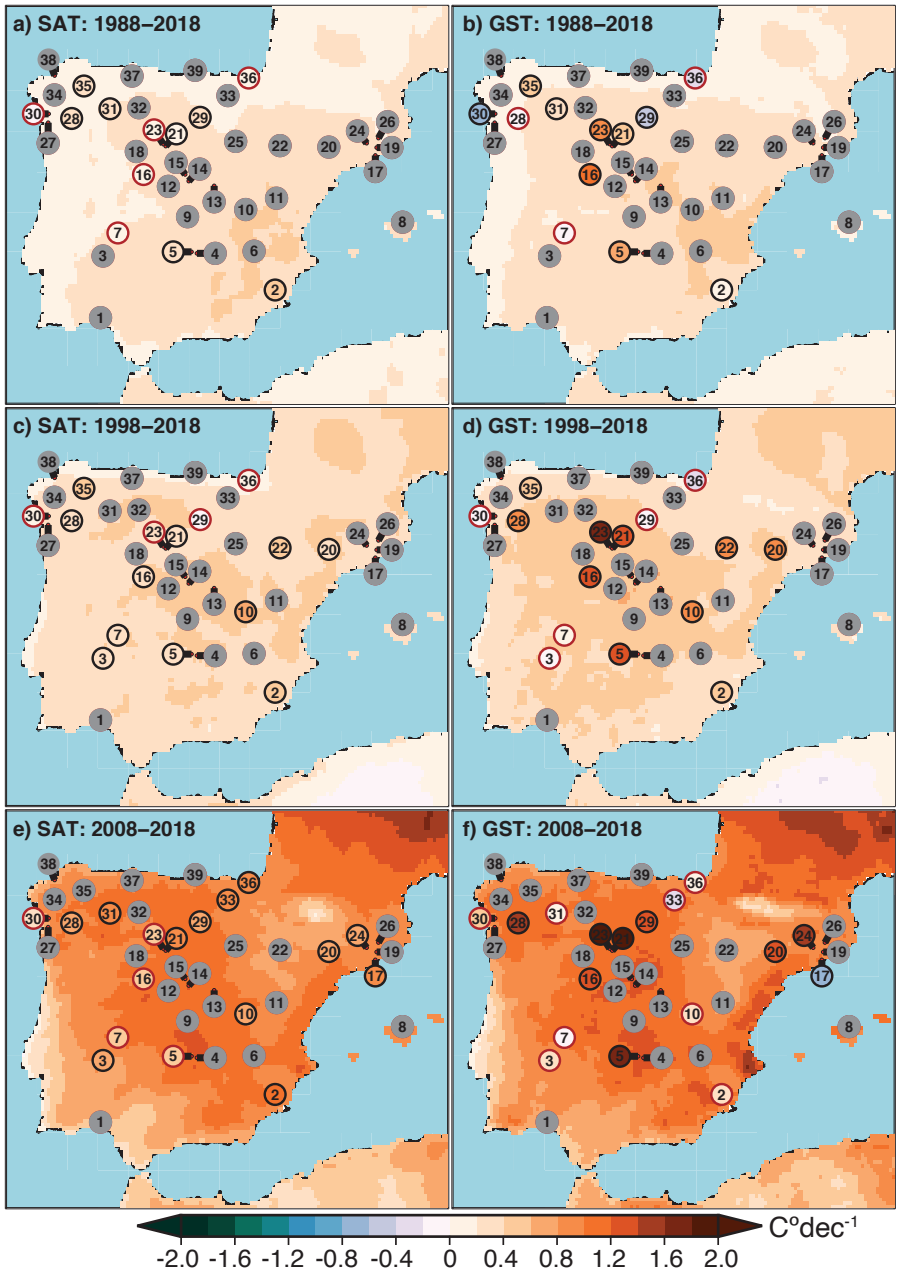

Figure 3.13: Temperature trends for the 1988-, 1998- and 2008-2018 periods (top, middle and bottom, respectively) for SAT (left), and ST_{005} (right). Note that for each reference period the number of sites is different as a result of the data availability. Trends are calculated from monthly anomalies and indicated in °C decade^{-1}. Black (red) contour circles indicate statistical significant (non-significant) values ($p < 0.05$). The grey-colored sites represent those locations with no sufficient data to estimate the trends. The results from the ERA5-L dataset are shown in the background.

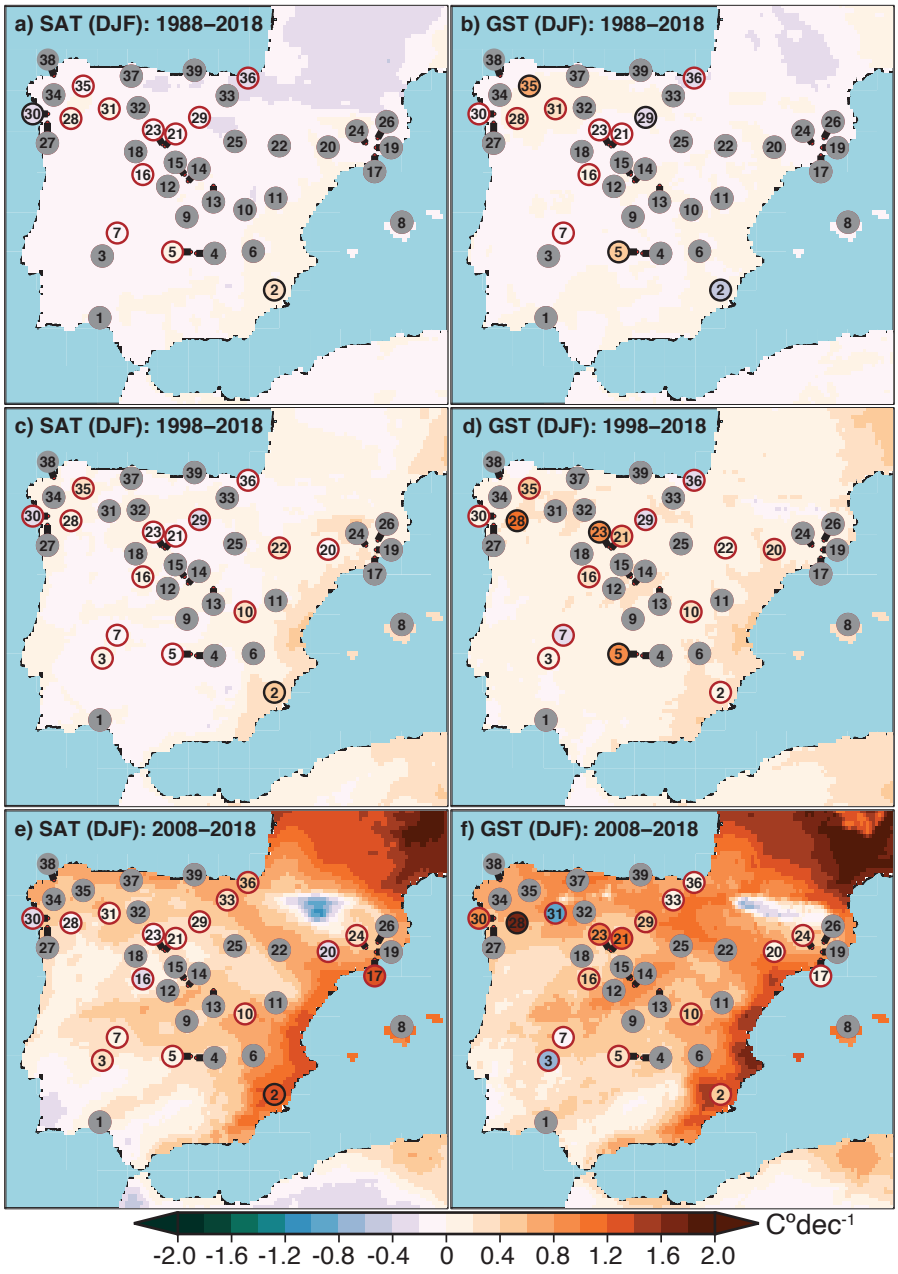

Figure 3.14: As Figure 3.13 but for DJF.

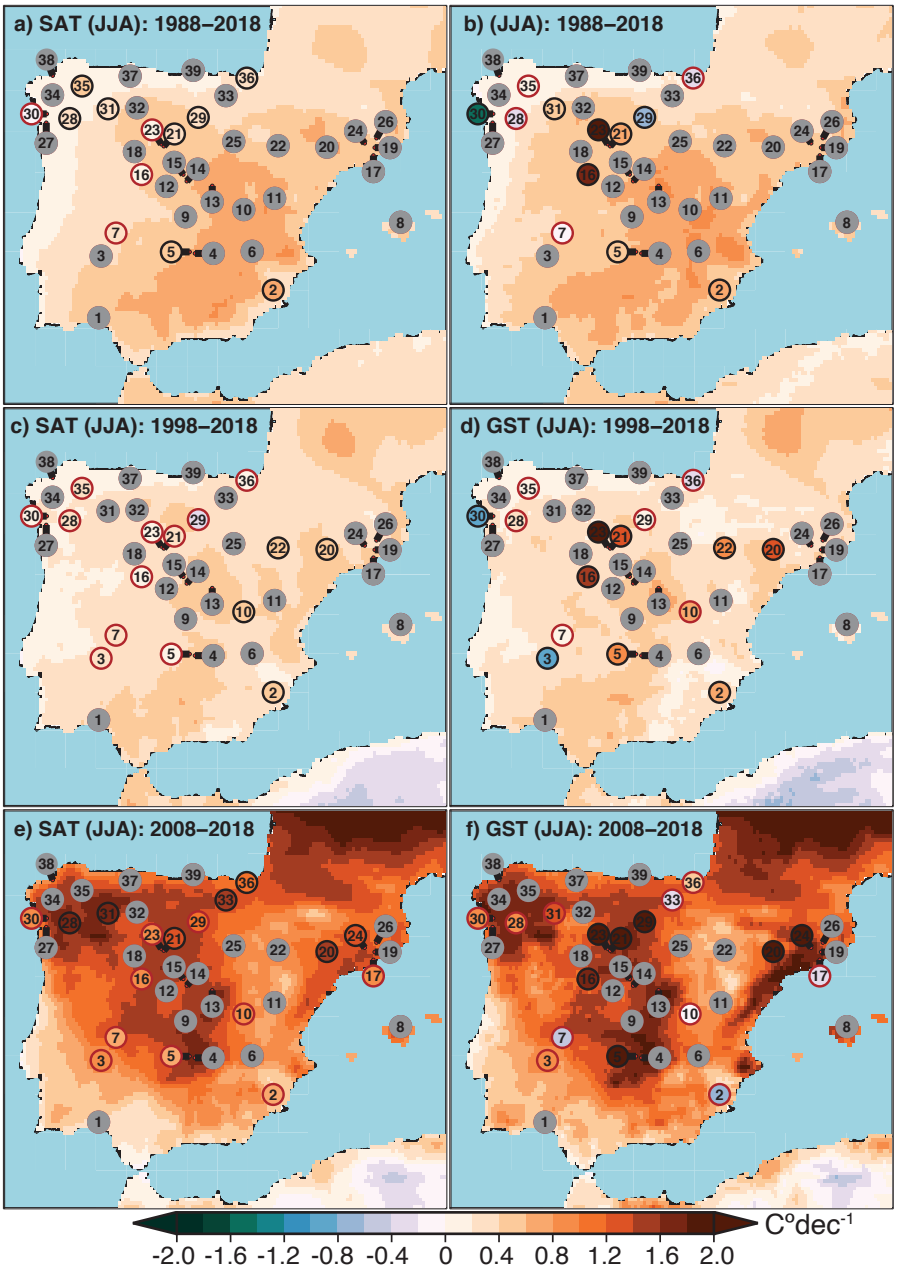

Figure 3.15: As Figure 3.13 but for JJA.

negligible positive trends in the 30- and 20-yr periods for both SAT and GST in the SoTOS and ERA5-L data, the JJA months show widespread warming over the Peninsula, denoting a larger contribution to the annual averages. Likewise, in JJA, the same pattern as the annual case of larger GST warming relative to the SAT for some SoTOS sites is observed, particularly over the Inner Plateau. Within the last decade, there is large warming of both DJF and JJA over the Mediterranean coast, which is also evident in the annual case. Interestingly, in DJF, ERA5-L shows a cooling trend over the Pyrenees for SAT and GST (Figure 3.14a,f) that imprints its signal in the annual average (Figure 3.13e,f). It is worth noting that during the last decade, the ERA5-L highly agrees with the SoTOS in representing the largely positive trends over the Inner Plateau and the Mediterranean coast, including the higher increase of GST relative to the SAT in these regions.

The evolution of SAT and GST shows some discrepancies in the annual and seasonal data for some of the SoTOS sites, with different magnitudes depending on the time frame considered. The largest contribution to the differences in the SAT and GST evolution comes from the JJA season. This agrees with the mean difference and correlation patterns described in Figure 3.10. The existence of changes in the SAT–GST relationship with time suggests some SAT–GST decoupling at interannual to multi-decadal timescales. Nevertheless, these results should be interpreted with caution as trends based on short records are very sensitive to the beginning and end dates due to the natural variability and generally do not indicate an actual long-term trend (Hartmann et al., 2013). Despite this, some consistent patterns hint at some decadal variations in the SAT–GST relationship over Spain. For instance, the negative trends over the northern Inner Plateau present in the three different periods of reference suggest that the higher warming rate of GST relative to the SAT (see Figure 3.13) is indeed the result of some changes in the surface energy balance at multi-decadal timescales.

Variations in the soil moisture content may explain the larger GST warming relative to SAT over the last few decades, especially in the summer months. As less water is available in the soil to sustain evaporation, the incoming solar radiation warms the ground surface by radiative heating. Therefore, GST can experience a larger temperature increase than SAT on average (Melo-Aguilar et al., 2018). Gallego-Elvira et al. (2016) evaluated the soil moisture control on land surface energy balance in an observation-based diagnosis using satellite observations and reanalysis data. They found that the land surface warms faster than the overlying atmosphere under water-stressed conditions due to limitations on evaporation and enhanced sensible heat fluxes. Additionally, they showed that this effect is larger over short vegetation and bare soil areas than forested regions due to differences in aerodynamic resistance and hydrological dry spell behavior.

Most of the SoTOS sites are located over bare soil and short grass surfaces (Figure 3.11a), thus enhancing the GST warming relative to the SAT effect. Likewise, a decrease in summer precipitation has been reported over the Iberian Peninsula during the recent decades (Río et al., 2011) as well as a decrease in surface humidity (Vicente-Serrano et al., 2014). Therefore, a reduction in soil moisture can also be expected. Addressing the specific causes of the variations in the SAT–GST relationship at these timescales would be desirable. However, this would require analyzing other meteorological variables to explore the physical mechanisms behind such responses in the SAT and GST. Such a level of information is not available in the present study, but it is highly interesting to develop such analyses in future works.

3.3 Conclusions

In the present chapter, an observational ST dataset for Spain has been employed to evaluate the conductive subsurface heat transport within the shallow subsurface and the strong coupling between SAT and GST. These two characteristics permit using present-day subsurface temperature measurements to retrieve the past surface temperature variations. Firstly, the one-dimensional steady-state analytical solution to the heat diffusion equation has been employed to evaluate the shallow subsurface thermal regime. This is achieved by assessing the conductive propagation into the soil of the surface temperature signal. The linear relationship of either the natural logarithm of the amplitude or the phase shift with the depth permits estimating the subsurface mean thermal diffusivity, assuming a purely conductive regime. The amplitude and phase of the ST at all depths have been extracted by a linear least-squares fit to a sinusoidal function of annual period for the 39 sites included in SoTOS. This information has been employed to estimate the soil mean thermal diffusivity over Spain as a measure to evaluate the thermal regime in the shallow subsurface. Secondly, the SAT–GST coupling assumption has been addressed by evaluating the SAT and GST covariance structure over decadal to multi-decadal timescales. In the present study, the ST at 0.05 m depth is considered as the GST as it is the closest soil layer to the ground surface. The SAT–GST coupling analysis is limited by the availability of ST data as not all of the 39 sites include a homogeneous temporal distribution of ST records for the complete 1985–2018 period. Therefore, only the sites with the longest records in ST data at 0.05 m depth are employed in this part. Despite this, a relatively good spatial distribution over the Iberian Peninsula and the Balearic Islands is

achieved, and in spite of its limitations, this is the only dataset we know of with this spatial and temporal extension over Europe.

This chapter also compares the observed air and soil temperature at all depths to the evolution of these variables, as represented by the current generation of land-surface reanalysis datasets. Such a comparison has provided useful information regarding the consistency between observed and land reanalysis datasets. Moreover, the reanalysis dataset has provided insights into the physical processes that influence both the thermal regime in the shallow subsurface and the SAT–GST coupling over Spain.

The results show that the conductive model explains most of the thermal regime within the shallow subsurface. Nonetheless, there are some deviations from the purely conductive propagation of the annual surface temperature signal. Such deviations translate to different results in estimating the mean thermal diffusivity of the shallow soil from the amplitude attenuation or the phase shift with depth for the bulk of the sites included in SoTOS. Similar works of this type have shown that using either the linear dependency of the amplitude attenuation or phase shift with depth should yield a similar estimation of the mean thermal diffusivity. However, in the shallow subsurface, some processes such as latent heat release/uptake may alter the conductive regime. The ERA5-L data has shown that the spatial distribution of water in soils influences the spatial variability in the soil thermal regime. Moreover, ERA5-L has also shown that variations in soil water content with time affect the thermal properties of soils since there is a high correlation between the variations in the soil thermal diffusivity and changes in the volumetric water content. Therefore, hydrology-related processes play an important role in the shallow subsurface's thermal regime that can alter the pure conductive heat transport assumption.

All in all, the results confirm that the heat transport within the soil is predominantly conductive despite the influence of hydrology-related processes within the shallow soil. The latter is supported by comparing the observed annual cycle of ST to the estimated one using the one-dimensional solution to the heat diffusion equation using either the estimated thermal diffusivity from either the amplitude attenuation or the phase shift. In both cases, the differences are minimal, thus supporting a predominantly conductive heat transport assumption.

Regarding the SAT–GST coupling assumption, the findings of this chapter indicate a strong connection between SAT and GST over Spain. Nevertheless, the SAT–GST bond is stronger in DJF compared to JJA in some of the SoTOS sites. During the summer months, the ground surface radiative heating leads to a relatively large SAT–GST offset with a general effect of warmer GST relative to the SAT. This physical response is mainly driven by the high incoming solar radiation in JJA over most of the Iberian Peninsula and the low precipitation

rates within these months, resulting in dry ground surface conditions. As water becomes a limiting factor, the incoming energy heats the ground surface faster than the air above. This is a general pattern over the center and the southern part of the country. The predominant location over bare soil and short grass surface characteristics of the sites included in SoTOS enhance the ground surface's radiative heating during the summer months. Some consistent variations in the SAT–GST relation at interannual to multi-decadal timescales over some regions of Peninsular Spain, specifically over the Inner Plateau, have been found in the SoTOS data. This is explained by a higher temperature increase of the ground surface compared to the SAT over the last three decades. Therefore, some SAT–GST decoupling may exist at these timescales. The ERA5-L, on the other hand, does not show the same SAT–GST decoupling as it underestimates the decadal ST trends over the 1985–2018 period. The nature of SAT–GST decoupling in SoTOS can be found in the decreasing precipitation and surface humidity trends in JJA during the last few decades over the center of the country, thus leading to a larger rate of sensible heat fluxes than latent heat dissipation.

Even though the results suggest that some SAT–GST decoupling may exist at decadal timescales, this does not necessarily imply that the SAT–GST coupling is corrupted at longer timescales. Similar assessments should be developed. The use of ST observational datasets with longer temporal coverage would be desirable. However, the availability of ST measurements is limited both in space and time. Therefore, the use of ESM output emerges as an opportunity to explore these issues for longer timescales.

4

Influence of radiative forcing factors on the SAT–GST relationship during the LM*

Chapter 3 focused on analyzing the thermal regime in the shallow subsurface in a case study using the SoTOS dataset. Due to the limited time span of the dataset (1850–2018), the relationship between SAT and GST could address interannual to multi-decadal timescales. Therefore, a longer time frame is required to address the long-term SAT–GST coupling and the mechanisms that may influence it. For this purpose, model simulations are required since observational records, especially for ST data, are limited both in space and time.

This chapter addresses the relationship between SAT and GST at yearly to multi-centennial timescales. This is initially done from a global perspective, but also a detailed analysis of the SAT and GTS responses at regional and local spatial domains is provided. The millennial-length model output from the CESM–LME (see Section 2.3 for a detailed description) is used for this purpose. As stated in Chapter 1.1, the use of LM simulations to explore the long-term SAT–GST relationship has shown that the SAT–GST coupling is indeed strong at global scales and above centennial timescales (e.g., González-Rouco et al., 2003, 2009; García-García et al., 2016). Nevertheless, some of the LM simulations employed

* The main contents of the chapter are included in the published research article:
- Melo-Aguilar, Camilo, Jesús Fidel González-Rouco, Elena García-Bustamante, Jorge Navarro and Norman Steinert, 2018: Influence of radiative forcing factors on ground–air temperature coupling during the last millennium: implications for borehole climatology. *Climate of the Past*, **14**, 1583–1606. DOI 10.5194/cp-14-1583-2018.

in the previous works lack full representation of the whole set of LM external forcings, and some others do not allow for the assessment of the individual contribution of individual external forcings, which is desirable in order to identify the nature of SAT–GST decoupling processes if any. Therefore, this chapter expands on previous assessments by addressing the relationship between SAT and GST during the LM for a range of spatial and temporal scales and its relationship with various mechanisms that impinge on it. Section 4.1 describes the methodological approach implemented in this chapter. Section 4.2 presents a general description of the covariance structure between SAT and GST at a global scale and the spatial distribution of the SAT–GST relationship over the LM via correlation and differences patterns. Further, a detailed description of the physical mechanisms explaining the cases in which SAT and GST exhibit different responses at annual or seasonal domains is provided. Finally, Section 4.2.1 explores the land-atmosphere interactions controlling the long–term SAT–GST decoupling processes during the LM, including the specific role of individual external forcings.

4.1 Methods

In the present chapter, LM refers to the period from 850 to 2005 CE while the periods from 850 to 1850 and 1851 to 2005 CE refer to pre-industrial and industrial, respectively. Two-meter air temperature is used for the SAT, and the first land model level (Table 2.4) represents GST. Additionally, the ST at different depths, denoted by the subscript notation: ST_{L2}, ST_{L3}, ..., ST_{L15}, is used to address different aspects of the subsurface heat transport within the CESM-LME. The simulated ST does not include data over the Antarctic region; therefore, it is excluded from the analysis.

The relationship between SAT and GST is analyzed from two complementary perspectives. First, the covariance during the LM of the global SAT and GST anomalies relative to 1851-2005 CE is assessed at an interannual scale. Likewise, the spatial patterns of the correlation between SAT and GST are also analyzed. This spatial analysis is extended to the differences between SAT and GST LM mean values as this gives an additional measure of the energy exchange across the air-ground interface (Bartlett et al., 2004) that may help characterize a potential air-ground temperature decoupling. The same spatial analysis is carried out for the differences between GST and ST_{L8} ($\sim$1 m depth) to get insights into processes affecting the pure conductive subsurface transport assumption within the shallow subsurface. For specific cases where SAT and GST exhibit some de-

coupling, a description of the main processes that lead to this decoupling at local and regional scales is presented. For this purpose, the evolution of SAT, GST and ST at different depths is assessed during the last 105 years of the simulation period (1900–2005 CE) for interest cases. In some particular examples, additional variables (e.g., latent and sensible heat fluxes or snow cover) are also included to describe the physical mechanisms better. The purpose of such analyses is providing some examples rather than detailing the minutiae of the processes that lead to different SAT–GST responses at many different locations. For the analysis described above, the ALL-F ensemble will be used. This provides a more complete and realistic representation of real-world conditions than the single forcing runs, from the point of view of the forcings considered. It also allows for the use of the surrogate reality of the CEMS-LME model as a test bed for detecting potential sources of deviations in the SAT–GST relationship during the LM.

This chapter is then focus on the assessment of the long-term trend in the SAT–GST relationship. Non-stationarity in such association at long timescales (multi-decadal to multi-centennial) might imply that GST variations would not represent the SAT variations; therefore, this would result in unreliable inferences of past climate change (Bartlett et al., 2004). A two-phase regression model (Solow, 1987) is applied to address this issue, which allows for detecting changes in the long-term trend of the SAT–GST relationship without imposing any temporal *a prior* condition on their occurrence. The year-of-change is identified, as well as the trend before and after the change. The timing of change and the magnitude of trends are suggestive of long-term changes in the surface coupling. Initially, this assessment is performed on a global scale by analyzing the LM evolution of global SAT minus GST anomalies. The long-term trend analysis is then extended, starting from the evaluation of SAT and GST linear trends independently during the industrial period since the largest changes are expected to occur within this period. In this part, besides the ALL-F ensemble, it is also considered the GHG-only, LULC-only and OZ/AER-only ensembles because they have the highest potential of altering the land surface characteristics. This allows identification of the influence of each forcing on the LM SAT–GST coupling, including changes in snow cover, soil moisture, latent and sensible heat variations. The linear trend analysis of SAT and GST gives a general view of their evolution through industrial times and a first glance at the most important areas where long-term SAT–GST decoupling may exist in response to changes in external forcings and related feedbacks. Although these processes have mostly taken place during industrial times, GHGs and land use have also experienced notable changes before 1850 CE (Ramankutty and Foley, 1992; Pongratz et al., 2008; Otto-Bliesner et al., 2016). The application of the two-phase regression to the complete LM permits identifying changes that occur before 1850 CE.

4.2 Results

The CESM-LME supports the assumption that the SAT is tightly coupled with GST at global scales and longer than multi-decadal scales. Figure 4.1a illustrates the stability of the coupling at global scales during the LM using the ALL-F_2 ensemble member as an example. Results are comparable for other ALL-F ensemble members. The LM evolution of global continental SAT, GST, ST_{L8} and STL_{L15} anomalies relative to the 1850–2005 mean are shown. For ST_{L15}, non-filtered model output is represented and evidences the low-pass filter influence of the heat conduction below the surface, whereas, for SAT, GST and ST_{L8}, 31-year running mean low-pass filter outputs are shown. Subsurface temperature anomalies closely track SAT anomalies with relatively small differences between them; this indicates that air and soil temperatures are coupled above multi-decadal timescales. The correlation coefficients for the 31-year filtered series in Table 4.1 indicate a high correlation ($p<0.05$) for the soil layers close to the surface that diminishes slightly with depth, as expected, due to the phase shift of the signal. Table 4.1 also indicates that the correlation considering the high-frequency variations (yearly; left column) is only high at the levels close to the surface, whereas, at the deepest layer, the correlation is low since the high-frequency variations are progressively filtered out and phase shifted as depth increases. Hence, the assumption of conductive heat transfer within the subsurface is realistically represented in the CESM-LME.

Despite the strong coupling between air and subsurface temperatures at global scales, the existence of a relatively small offset between SAT and GST that grows backward in time is evident for the annual averages (Figure 4.1a). This indicates a slight long-term decoupling between SAT and GST. Previous works have argued that the nature of this offset arises from the changes in snow cover and its influence insulating the ground from freezing air temperatures (e.g. Mann and Schmidt, 2003; Bartlett et al., 2004). To explore the influence of such processes on the global SAT–GST relationship, this analysis is extended considering DJF and JJA independently (Figure 4.1b,c). The offset between SAT and GST observed in the annual plot is apparent only in the DJF season. Indeed, the differences are larger during this season than in the annual data, whereas the JJA evolution of SAT and GST anomalies is virtually identical. Even if SAT and GST are highly correlated and significant for both seasons, Table 4.1 suggests a slightly lower correlation for DJF than for JJA. This feature and the largest offset occurring in DJF suggest an important role by snow cover.

Figure 4.2 further illustrate the strong relation between the SAT–GST offset and snow cover by displaying the 31-year low-pass filter outputs of the LM evolution of global snow cover and SAT–GST differences. Note that in DJF (Figure

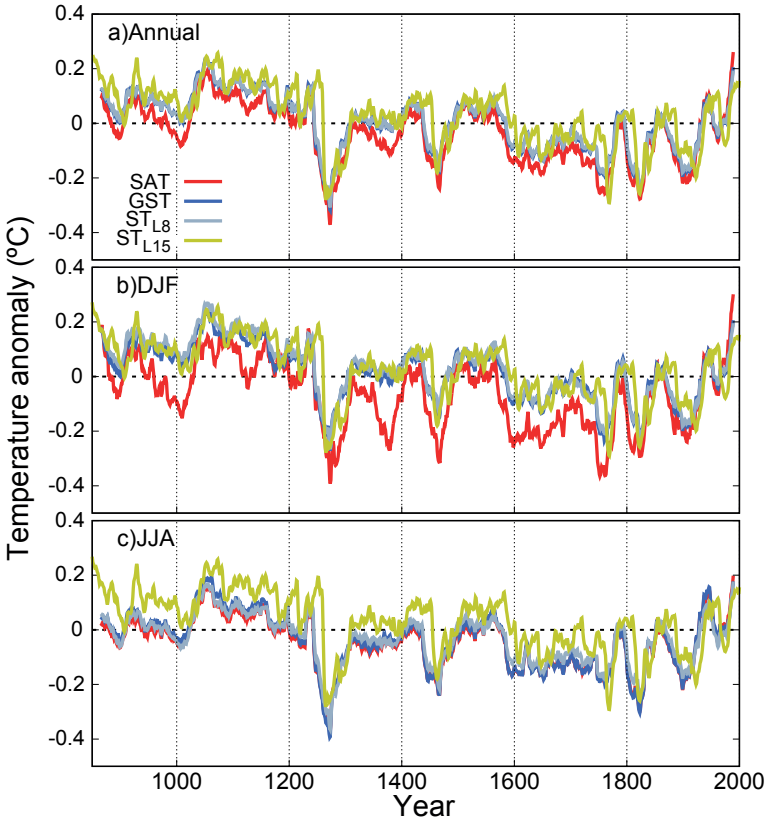

Figure 4.1: (a) Annual, (b) DJF and (c) JJA global LM evolution of SAT, GST, ST_{L8} and ST_{L15} anomalies relative to 1850-2005 CE for the ALL-F_2 ensemble member. All series are 31-yr moving average filter outputs except for ST_{L15} is represented as the direct model output.

4.2b), due to its influence on annual averages, the decrease (increase) in snow cover leads to a decrease (increase) in the SAT–GST offset in absolute values. This is due to the insulating effect of snow that keeps GST close to zero while SAT can reach large negative values. Thus, an increase in snow cover leads to larger negative SAT–GST differences. For JJA on the contrary, an in-phase relationship is found at all timescales. Long-term trends change in both snow cover and in SAT–GST after the end of the 19th century. During the boreal summer

Table 4.1: Temporal correlation coefficients between last millennium SAT–GST, SAT-ST_{L8} and SAT-ST_{L15} anomalies relative to the 1850-2005 mean for the experiment shown in Figure 4.1. The left side indicates the correlation at yearly resolution whereas the right side shows 31-yr low-pass filter outputs of SAT, GST, ST_{L8} and ST_{L15} for annual, DJF and JJA. Coefficients highlighted in bold are statistically significant at $p<0.05$

	Yearly			31-yr filtered		
	Annual	DJF	JJA	Annual	DJF	JJA
SAT–GST	**0.96**	**0.85**	**0.99**	**0.98**	**0.94**	**0.99**
SAT-ST_{L8}	**0.95**	**0.81**	**0.94**	**0.98**	**0.93**	**0.96**
SAT-ST_{L15}	0.28	0.27	0.14	**0.86**	**0.85**	**0.65**

increases in snow also enhance SAT–GST differences due to the insulation of the ground from the warmer summer SAT, whilst the opposite is noted for snow cover decreases. This effect dominates the global average over that of the JJA austral winter during which SAT–GST and snow cover changes experience an anti-phase relationship as described above. Therefore, the NH influences anti-phase covariability of snow cover and SAT–GST during DJF (detrended correlations, r=-0.52; $p<0.05$) and annual (r=-0.67; $p<0.05$) and in-phase covariability during JJA (r=0.62; $p<0.05$).

Beyond the anti- and in-phase covariability at multi-decadal to centennial timescales, changes in the longer-term relationship between SAT and GST can play an important role in decoupling with implications for the borehole theory. At global scales, the long-term offset is relatively small, as shown in Figure 4.1, and therefore has minimal implications. Nevertheless, it is interesting to assess the consistency and relative magnitude of snow cover and SAT–GST changes. Two-phase regression in annual and seasonal SAT–GST (Figure 4.2) shows a consistent change date during the 18th century. All changes are towards smaller SAT–GST differences and are significant for annual and DJF ($p<0.05$). For snow cover trends, using the same dates of change as for SAT–GST, indicate anti-phase (in-phase) relationships for the annual and DJF (JJA) periods as in the multi-decadal timescales described above. However, changes in snow cover trends during the last centuries are small and can hardly be invoked to solely account for the comparatively larger SAT–GST trend changes. This calls for the consideration of other possible mechanisms and a more spatial perspective.

The spatial variability of the relationship between SAT and GST gives further insights into the role of different processes on the SAT–GST coupling. Figure 4.3

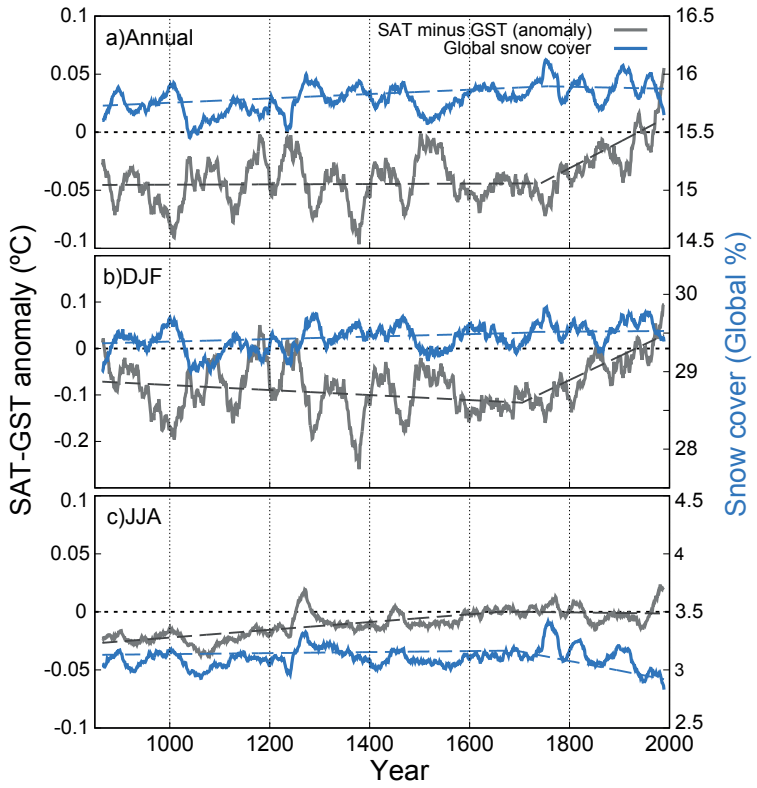

Figure 4.2: (a) Annual, (b) DJF and (c) JJA global LM evolution of SAT–GST offset (SAT minus GST anomalies) and global percentage of snow cover for the ALL-F_2. Dashed lines are the result of a two-phase regression, indicating the linear trend that represents the best fit to the data before and after the estimated point of change. For snow cover, dashed lines represent linear fits to the data using the change points found for SAT–GST. All series are 31-yr moving average filter outputs.

shows the spatial distribution of the differences between mean LM SAT and GST as well as the correlation coefficients for the annual, DJF and JJA averages in the ALL-F ensemble. In the case of the annual temperatures (Figure 4.3a), GST is generally warmer than SAT, with the differences being low over most of the globe (less than 2 °C) except in the NH mid- and high-latitude areas where the

differences are higher (up to 15 °C). The correlation maps (Figure 4.3b) provide a similar pattern with high and significant values over most of the globe (>0.8 in regions located between 45° N and 90° S) and lower correlations over NH mid- and high-latitudes, especially over eastern Siberia. Similar behavior is seen in the DJF season, although it is more pronounced. During these months, GST is much warmer than SAT, reaching differences up to 30 °C in the northernmost parts of North America and Eurasia (Figure 4.3c). Differences are smaller at mid- and low-latitudes. The influence of the ocean over coastal areas providing larger SAT relative to GST is noticeable. Similarly, the correlation is lower over the northern snow-covered areas (Figure 4.3d), while over the rest of the globe, it remains high. In contrast, Figure 4.3e,f show that during JJA, when the snow cover is scarce, the SAT–GST coupling is strong globally with temperature differences lower than 2 °C and high correlation coefficients (>0.9). Consequently, the role of snow cover in decoupling SAT and GST is highlighted. Positive correlation values are low over borderline areas where snow cover is more variable, which produces variability in the SAT–GST offset and thereby alters the covariance structure. Close to these areas in central Asia, the Tibetan Plateau's high negative correlation is noteworthy (see discussion below).

Figure 4.4 illustrates the SAT–GST decoupling due to the snow cover at the local scale for a particular grid point as an example. The grid point is located over a region with considerable snow cover during the boreal cold season (northeastern Russia). The DJF and JJA evolutions of SAT, GST and ST at different depths and the snow cover for the last 105 years from the ALL-F_2 simulation are shown. Note that in DJF, snow covers 100% of the grid cell during almost the whole period. Thus, the soil is insulated, and the difference between SAT and GST is, on average, ∼-15 °C. The temperature of the deeper layers is presented to illustrate the amplitude attenuation and phase shift with the depth of the temperature signal. Note that during DJF, GST is only slightly below 0 °C, and the agreement of its variations with those of SAT is only noticeable in the largest changes of both, while ST_{L6} and deeper STs are above 0 °C. In JJA, SAT and GST are very similar and their low frequency variability propagates to deeper levels, all above 0 °C.

Some aspects of the SAT–GST spatial distribution deserve further attention. One of the most noteworthy of these is that over the Tibetan Plateau region, temperature differences between SAT and GST are as large as in the mid- and high-NH latitudes for both annual and DJF periods (Figure 4.3a, c). However, SAT and GST are negatively correlated for annual and DJF seasonal resolution data (Figure 4.3b,d), with the Tibetan Plateau being the only region of the globe where this occurs. Nevertheless, the correlation for JJA is positive and high (Figure 4.3e). The nature of this opposite phase arises from discontinuous snow cover

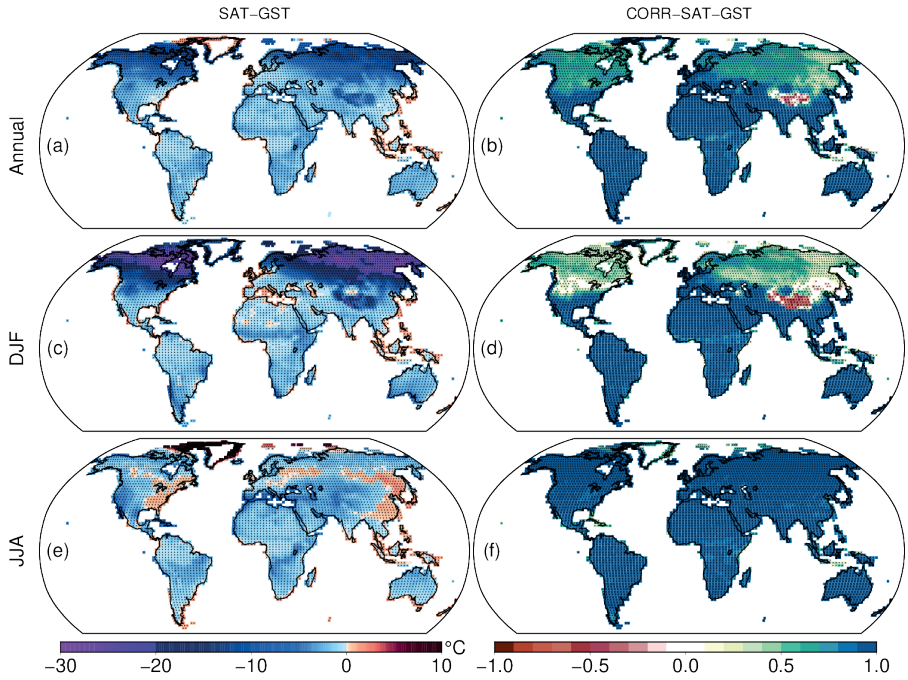

Figure 4.3: Left. (a) Annual, (c) DJF, (e) JJA spatial distribution of SAT minus GST differences during the 850–2005 CE period. Right. (b) Annual, (d) DJF, (f) JJA spatial distribution of the correlation coefficients between SAT and GST for the same period. Black dots indicate that 80% of the members within the ALL-F ensemble deliver statistical significance (p<0.05).

over this region and the very low air temperatures during DJF. Usually, the snow cover insulates the soil from the colder air, avoiding the heat exchange with the atmosphere (as shown in Figure 4.4). Nevertheless, discontinuous snow cover only partially insulates the soil, leading to this particular SAT–GST interaction. Figure 4.5 displays this behavior for a grid point located over the Tibetan Plateau. SAT, GST, ST at different depths, snow cover and surface sensible heat flux (SHFLX) are shown. In DJF, during periods of low snow cover, the fraction of surface exposed to the atmosphere allows for energy exchange from the warmer soil to the colder air. Conversely, when snow cover is high, the large fraction of insulated soil reduces almost completely the heat transfer from the soil to the at-

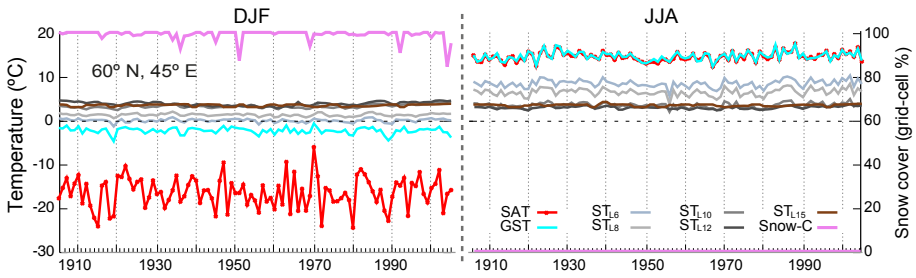

Figure 4.4: Evolution of DJF and JJA SAT, GST, ST_{L6}, ST_{L8}, ST_{L10}, ST_{L12} and ST_{L15} and percentage of snow cover during the period 1900-2005 CE for a grid point from the ALL-F_2 simulation (northeast of Russia; 60°N,45°E) where snow cover is characteristic during the cold season. Dashed lines indicate 0 °C.

mosphere. Therefore, with lower (higher) fractions of snow cover, higher (lower) heat transfer takes place with GST decreasing (increasing) and SAT increasing (decreasing). Indeed, there is a high negative correlation (-0.84; $p<0.05$) between the snow cover fraction and SHFLX, which is the main way that energy dissipates within this region since latent heat fluxes (LHFLX) are negligible. Higher/lower albedo due to variations in snow cover fraction also contribute to the negative SAT–GST correlation over this region (not shown). The horizontal model resolution does not seem to be an issue since a higher resolution version of the model (Landrum et al., 2013) produces a similar behavior which other models of similar resolution do not show (García-García et al., 2016). During JJA, in comparison, snow cover is negligible and SAT–GST coupling is consequently strong. Note that GST and all ST are above zero, and GST is higher than SAT as it is warmed by radiative gain and the transfer of heat to the atmosphere, hence the positive correlation(0.5; $p<0.05$) between SAT–GST and SHFLX.

The spatial SAT–GST differences during JJA (Figure 4.3e) depict other relevant aspects influencing the SAT–GST relationship at relatively short timescales. SAT is generally colder than GST globally. However, for JJA, there are large areas inland, mainly located in the southeastern US, some parts of central and eastern Europe and eastern Asia, with warmer SAT relative to GST. Variations in LHFLX from DJF to JJA drive this effect. Figure 4.6 shows that the areas where LHFLX increases in JJA relative to DJF are related to the same areas where the SAT is higher than GST in JJA (Figure 4.3e). Therefore, there is a direct relation between the increase in evapotranspiration in JJA and the ground temperature response at these locations. The time series in Figure 4.6 display

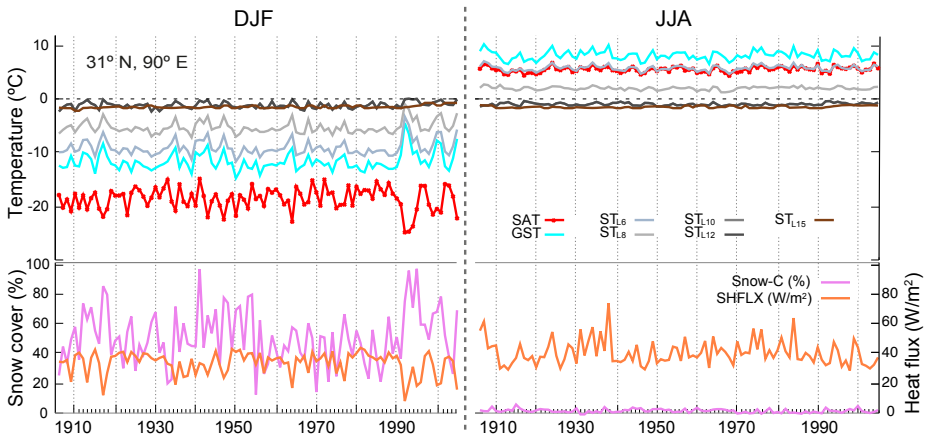

Figure 4.5: Top. DJF and JJA evolution of SAT, GST, ST_{L6}, ST_{L8}, ST_{L10}, ST_{L12} and ST_{L15} for a grid point from the ALL-F_2 simulation (Tibetan Plateau region; 31°N,90°E) during the period 1900–2005 CE. Dashed lines indicate 0 °C. Bottom. Percentage of snow cover and surface sensible heat flux (SHFLX).

this behavior for a grid point located over southeastern China. During JJA, the surplus of energy due to higher solar radiation reaching the surface is mostly dissipated as latent heat leading to a net heat loss at the ground surface. Note the anticorrelation between the surface soil moisture and LHFLX (-0.46; $p<0.05$) during JJA and the large SHFLX values that occur when soil moisture is at its lowest and evapotranspiration is limited. Therefore, the high rate of LHFLX contributes to cool the surface, and GST tends to be lower relative to SAT. Soil water content also exhibits large changes during JJA consistent with the large evapotranspiration (Figure 4.6, top) and provides a moisture source that contributes to temperate SAT and cool GST. Large variations in evapotranspiration from DJF to JJA are also present at midlatitudes of the SH summer continents (America, Africa and Australia), although only a minimal impact on Figure 4.3e is perceived, especially over the western coast. Over these regions, high incoming energy impinging the surface during SH summer (not shown) supports high latent heat flux rates. However, as soil water becomes a limiting factor, more energy is dissipated as sensible heat, and the ground surface is warmed. Therefore, GST experiences higher temperatures than SAT on average. Figure 4.6 also shows high evapotranspiration over the tropical rainforest in America and Africa, both in DJF and JJA, which does not translate to positive SAT–GST differences. Over

the rainforest, the energy fluxes at the surface do not vary significantly from DJF to JJA since incoming radiation is relatively constant throughout the year and is transferred to evaporation and evapotranspiration within the canopy. At the same time, soils are well watered by precipitation to support the large amounts of evapotranspiration. This situation leads to a small range of variation in both SAT and GST, with very small differences between them and higher GST relative to SAT.

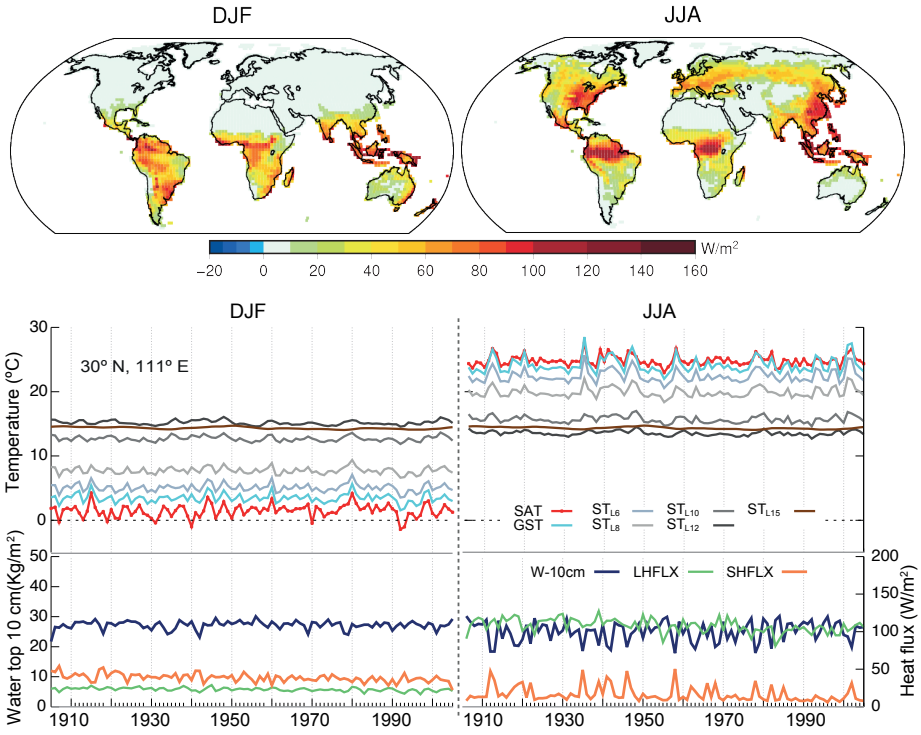

Figure 4.6: Top. DJF and JJA average surface latent heat fluxes (LHFLX) over the 850-2005 CE period. Middle. DJF and JJA time evolution of SAT, GST, ST_{L6}, ST_{L8}, ST_{L10}, ST_{L12} and ST_{L15} during the period 1900-2005 CE for a grid-point from the ALL-F_2 simulation (southeastern China; 31°N,111°E). Bottom. Same for SHFLX, LHFLX and water content in the top 10 cm of soil (W-10cm).

Figure 4.3a,c, and e show similar higher SAT values relative to GST at some coastal areas in JJA and DJF. During DJF, the effect is mainly present in the NH midlatitudes, such as in most European coastal areas, both the east and the west coasts of North America, and in Japan and the east coast of China. During JJA, in comparison, this behavior is mostly seen in the SH midlatitudes, such as southern South America, South Africa and southern Australia. Interestingly, higher SAT relative to GST is also evident in some coastal areas over tropical regions; this applies for both JJA and DJF and is mainly observed over southeastern Asia, the southern areas of the Indian subcontinent, the Gulf of Guinea and some areas of South and Central America. In the CESM-LME, the atmospheric grid box of the coastal areas is partitioned into land and ocean fractions. For sea ice formation areas, an additional sea ice fraction is considered (Neale et al., 2012). This configuration of the coastal grid points leads to a partition of the energy fluxes at the surface into those of the land fraction and the ocean fraction. During the cold season, partitioning such as this determines the higher SAT warming relative to GST, as the relatively low net radiation that impinges the surface at mid- and high-latitudes limits the ground surface heating and the energy fluxes out of the land surface fraction. In contrast, the energy fluxes from the ocean surface to the air above are large, primarily due to the temperature difference between the water and the comparatively colder air above. The dissipation of energy from the ocean fraction to the atmosphere warms the air, so the net effect is higher SAT relative to GST in the winter season. Over the tropical coasts that exhibit the same behavior, the energy fluxes out of the ocean fraction of each grid point also contribute to the higher warming of the air relative to the ground surface. Nevertheless, at these locations, high evapotranspiration rates all year long also play an important role as they generate evaporative cooling of the ground surface, as seen in the example described in Figure 4.6.

The different examples used to illustrate the most important processes that may influence the air and soil temperature relationship at short timescales also depict relevant information about the propagation with depth of the annual cycle. For instance, at the grid point located over southeast China in Figure 4.6, both in DJF and JJA, the temperature offset between contiguous levels is noticeable with a gradient of about 5 °C in the first meter of the ground and about 10 °C down to the lowest level. Comparable pictures with some differences in the magnitude of gradients can be seen in the previous figures.

Therefore, it is interesting also to understand the propagation of temperature below the surface. GCMs simulate purely conductive regimes, and the temperature variations that propagate to deeper soil layers are established at or near the ground surface (Smerdon et al., 2003). Thus it is important to assess the

propagation of the temperature signal within the shallow subsurface. This issue is addressed by analyzing the relationship between GST and ST_{L8}.

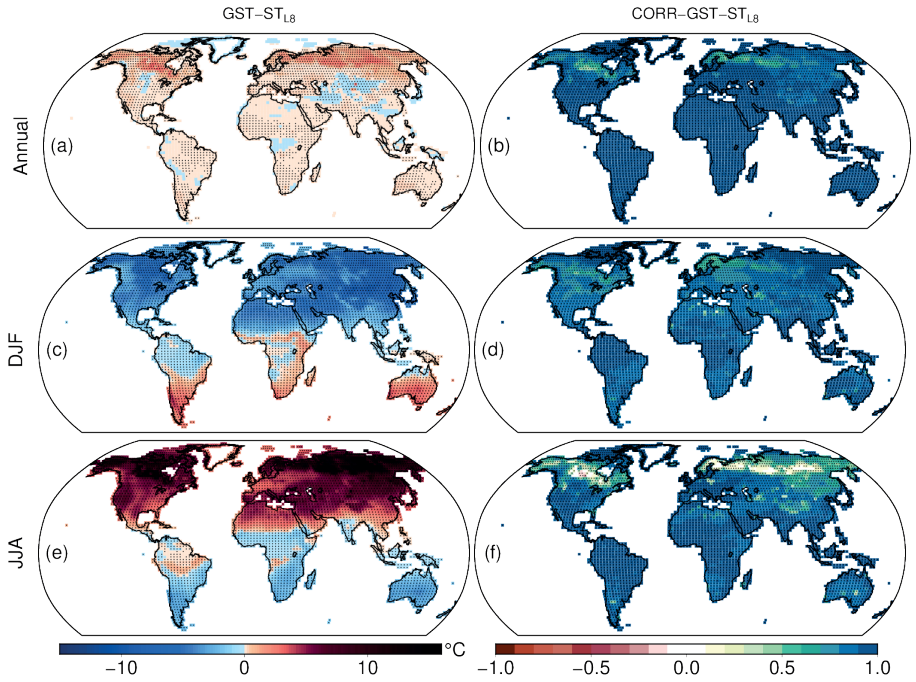

Figure 4.7: (a) Annual, (c) DJF, (e) JJA spatial distribution of GST minus ST_{L8} differences during the 850–2005 CE period. (b) Annual, (d) DJF, (f) JJA spatial distribution of the correlation coefficients between GST and ST_{L8} for the same period. Black dots indicate that 80% of the members within the ALL-F ensemble deliver statistical significance (p<0.05).

Figure 4.7 provides a spatial view of the temperature differences between GST and ST_{L8} for annual, DJF and JJA periods. The correlation is also shown in the right panels. SAT–GST differences for DJF and JJA show the yearly cycle of temperature with negative (positive) SAT–GST for the NH (SH) in DJF and vice-versa for JJA, illustrating the conductive regime within the shallow subsurface. The annual temperature differences are low, and the correlation is high almost globally as it is the balance between the respective patterns in JJA and DJF. However, the northernmost part of the globe exhibits larger temperature

differences (between 4 and 5°C) and lower correlation coefficients. The DJF and JJA patterns show that the annual offset and correlations for this part of the globe are mostly the result of the larger weight of those in JJA given that, during these months, the temperature differences for a latitudinal band at ca. 60-70° N are as large as 15 °C and the correlation coefficients are close to zero. García-García et al. (2016) describe a similar behavior in some of the GCMs used in their analysis, detected over areas where frozen ground persists during JJA. Indeed, the nature of the large departure in the temperature response at the shallow subsurface at these locations arises from non-conductive processes related to latent heat release/uptake of freezing and thawing of the water content above a depth of 1m that may account for the subsurface heat transfer (Kane. et al., 2001).

To illustrate this mechanism, Figure 4.8 shows the temperature evolution of SAT, GST and ST at different depths and the soil ice content (SIC) in the upper soil layers for a grid point located in the north of Canada. Over these areas, the SIC in the upper 1m of soil increases (decreases) during the cold (warm) season. During JJA, SAT and GST increase/decrease at the same rate since no ice is present at the ground surface, so it is warmed by radiative gain and heat is transferred to the atmosphere. However, for deeper soil layers, the energy available is employed to melt the SIC (note the lower SIC in L6 during JJA relative to DJF), and latent heat is required so that these layers do not experience a temperature increase like the shallowest layers. Therefore the temperature at L8 (∼1 m depth) is kept below/near 0 °C during the warm season due to the zero-curtain effect (Outcalt et al., 1990), while GST is centered around 12°C, leading to differences of ca. 15 °C between GST and ST_{L8} and a low correlation coefficient (0.28 for this grid point) during JJA. In turn, during DJF, the SAT sits around -35 °C and the frozen ground experiences skin temperatures of about -8 °C and about -4 °C at a depth of 1m. Consequently, the temperature offset between GST and ST_{L8} is largest in JJA. As a result, there are some non-conductive processes associated with permanently frozen soils in the shallow subsurface that are included in the CLM4 parameterization (Lawrence et al., 2011) and that play an important role in heat transport. The higher SIC in L8 during JJA relative to DJF (Figure 4.8) arises from the fact that freezing of the active layer begins in late autumn and, due to the release of the latent heat of fusion, the freezing front is inhibited (Hinkel et al., 2001) and only reaches L8 in spring when SIC at this layer peaks. Then, as the thawing front penetrates downward in JJA, ice in the shallowest soil is melted, but it does not reach L8 until autumn when SIC at this layer is lowest. Therefore, due to the thawing–freezing processes, seasonal changes at the upper and deeper subsurface levels are phase-shifted.

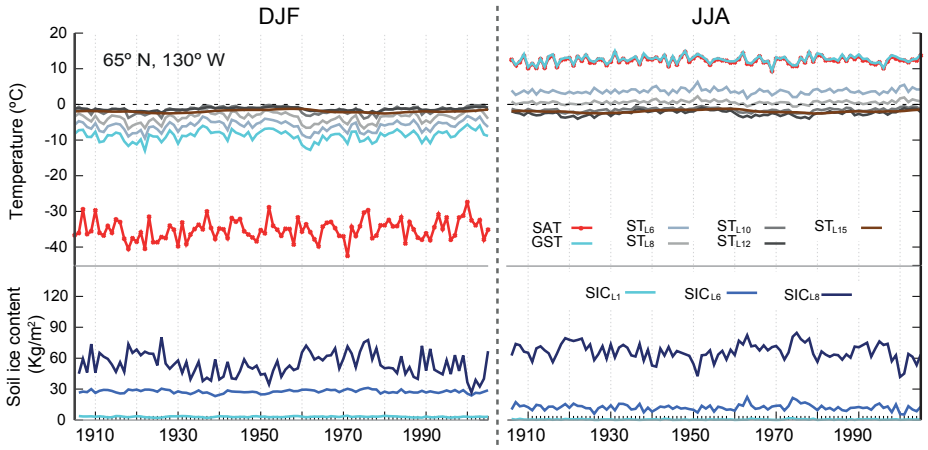

Figure 4.8: Top. DJF and JJA evolution of SAT, GST, ST_{L6}, ST_{L8}, ST_{L10}, ST_{L12} and ST_{L15} during the period 1900-2005 CE for a grid-point from the ALL-F_2 simulation located at the north of Canada (65°N,130°W), an area with permanent frozen soils. Dashed lines indicate 0 °C. Bottom. Soil ice mass content (SIC) in the L1, L6 and L8 soil layers.

4.2.1 SAT–GST long-term changes

The mechanisms that have been described impact the coupling between SAT and GST at short timescales, but they do not affect the long-term SAT–GST association, if they are stationary as its influence, would be constant at long timescales. However, if such mechanisms experience variations with time, the SAT–GST relationship would also change over time. Thus, the thermal signature imprinted in the subsurface would not be representative of the long-term SAT variations (Bartlett et al., 2005). Figure 4.2a illustrates the existence of a constant offset between SAT and GST within the preindustrial period that changes during the industrial period indicating variation in the long-term SAT–GST relationship. This may be relevant in interpreting borehole climate reconstructions because it may induce a long-term decoupling between SAT and GST in the CESM-LME. On a global scale, the long-term SAT-GST offset changes have an impact of about 0.05 °C (Figure 4.1a,b); thus, they do not seem to be very relevant at these scales. However, the impact could be larger for other GCMs with higher climate sensitivity or a different representation of surface processes that may contribute to decouple GST from SAT (e.g., snow cover). Similarly, within the

CESM-LME simulations, impacts on decoupling may be important at regional or local scales.

To examine the spatial distribution of the long-term SAT and GST evolution during industrial times, we evaluate the linear trends of both temperatures independently during this period at every land model grid point. Besides the ALL-F ensemble, this chapter also considers the anthropogenic single-F ensembles (Figure 4.9), bearing in mind their potential influence on the processes that modulate the relationship between SAT and GST, such as variations in snow cover, soil moisture and albedo, among others. The results in this section are shown considering information from all members of the ensembles (see details in Table 2.6). For specific examples, one of the members will be used as indicated accordingly in figure captions.

Figure 4.9a,b describes predominant warming for both SAT and GST in the ALL-F ensemble with the largest values distributed over northwest North America, north and central Eurasia, northeast Africa and southern South America. Interestingly, there are also regions showing negative trends like southeastern China, the north of the Black and Caspian sea regions, Pakistan, some relatively small central and southern areas of Africa and Brazil. The warming trend pattern can be explained to a large extent if the 1850–2005 trends are calculated based on the GHG-only ensemble (Figure 4.9c,d), which is consistent with the global warming pattern due to the influence of GHGs (Hartmann et al., 2013). Indeed, if only the contribution of GHGs is considered, the warming would be higher and globally distributed. Figure 4.9 also indicates that the cooling in the ALL-F ensemble is mainly driven by the LULC and OZ/AER external forcings' contribution. For instance, the cooling trends over the Baltic Sea and the north of the Black and Caspian seas that dominate the SAT and GST cooling trends during the industrial period in the ALL-F ensemble are the result of the influence of LULC changes (Figure 4.9e,f) with additional contributions of OZ/AER (Figure 4.9g,h). The negative trends of both SAT and GST over some areas of Africa and the northeast of Brazil are also detectable in the LULC-only ensemble (Figure 4.9e,f). Similarly, the OZ/AER-only ensemble also contributes to the cooling over Brazil, and the strong negative trends observed in southeast China in the ALL-F ensemble are clearly identifiable in this ensemble (Figure 4.9g,h).

Although the general pattern of cooling/warming during industrial times is broadly similar for SAT and GST, with a spatial pattern correlation of 0.60, 0.64, 0.38 and 0.53 in the ALL-F, GHG-only, LULC-only and OZ/AER-only ensembles, respectively, it differs substantially in some regions. Note that there are considerable differences in the amplitude of warming trends in SAT and GST over Fennoscandia and at the northernmost part of North America in the ALL-F ensemble (Figure 4.9a,b). Similarly, over some areas of central and eastern Europe,

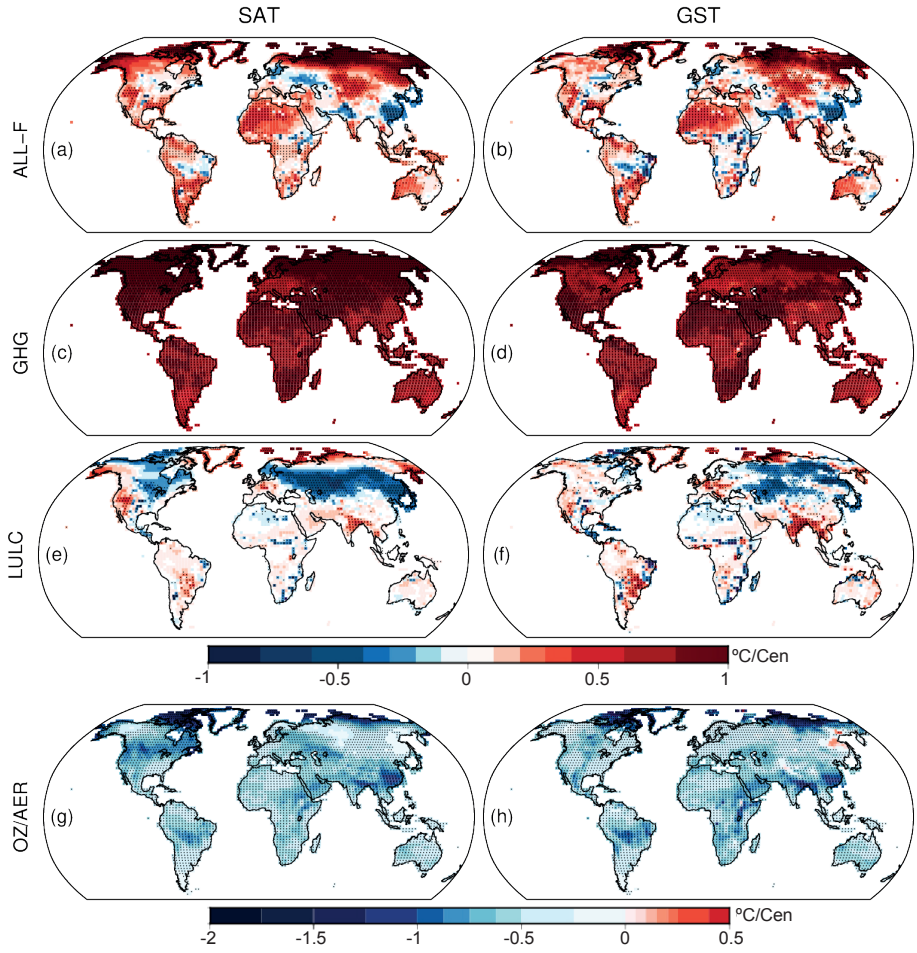

Figure 4.9: Spatial distribution of the linear trends in the industrial period for the ALL-F (a, b), GHG-only (c, d), LULC-only (e, f) and OZ/AER-only (g, h) ensembles for SAT (a, c, e, g) and GST (b, d, f, h). Trends are indicated in Celsius century^{-1}. Dots indicate that 80% of the ensemble members agree in delivering significant trends at a grid point in the case of the ALL-F ensemble. For the GHG-only and LULC-only ensembles, dots indicate that at least two of the three ensemble members agree in delivering significant trends for a grid point, whereas for the OZ/AER-only dots indicate that both of the ensemble members deliver significant trends. Note the different scale for OZ/AER-only.

SAT and GST industrial trends have different sign. There are also considerable differences in the amplitude of the cooling in SAT and GST over northeastern Brazil. Such dissimilar behaviors of SAT and GST during the industrial period are connected to variations in the energy fluxes at the surface in response to changes in the land surface characteristics due to the influence of the external forcings during this period.

For a more detailed analysis of the SAT and GST long-term relationship, a two-phase regression model (Section 4.1) was applied at every land model grid point to the SAT–GST differences in the ALL-*F*, GHG-only and LULC-only ensembles (Figure 4.10). This allows for the analysis of long-term changes in the coupling without assuming any a priori condition on their time of occurrence and separately analyzing the contribution of the different forcing factors and the temporal consistency among them. In the OZ/AER-only ensemble case, since this set of simulations spans from 1850 to 2005 CE, linear trends are only shown for the industrial period (Figure 4.11). Figure 4.10a shows the year of change for the three ensembles; changes significant for 80% of the ensemble members are shown exclusively in Figure 4.10b. The three ensembles show dates of change that span the whole millennium. However, significant changes only occur during the last centuries. Two-phase regression allows for identification of the fact that times of change in most regions occur before 1850, during the 18th and even the 17th century (e.g., India) in the ALL-*F*. Trends before the change are not significant in any of the ensembles (not shown). Trends after the change are shown in Figure 4.10c; significant areas in at least 80% of the ensemble members are exclusively shown in Figure 4.10d. Note that large positive and negative trends in Figure 4.10c coincide with the significant dates of change occurring during the last centuries of the LM.

In general, annual SAT minus GST yields negative values (as shown in Figure 4.3a) except for the coastal areas, as explained in Section 4.2. Thus, for continental areas (SAT–GST$<$0), positive (negative) trends indicate that differences tend to get smaller (larger) in absolute values, whilst the opposite is true for the limited coastal areas where SAT–GST differences are positive. Figure 4.2a allows for visualization of this behavior. Note the positive trend after the change when the difference between SAT and GST anomalies becomes smaller with time.

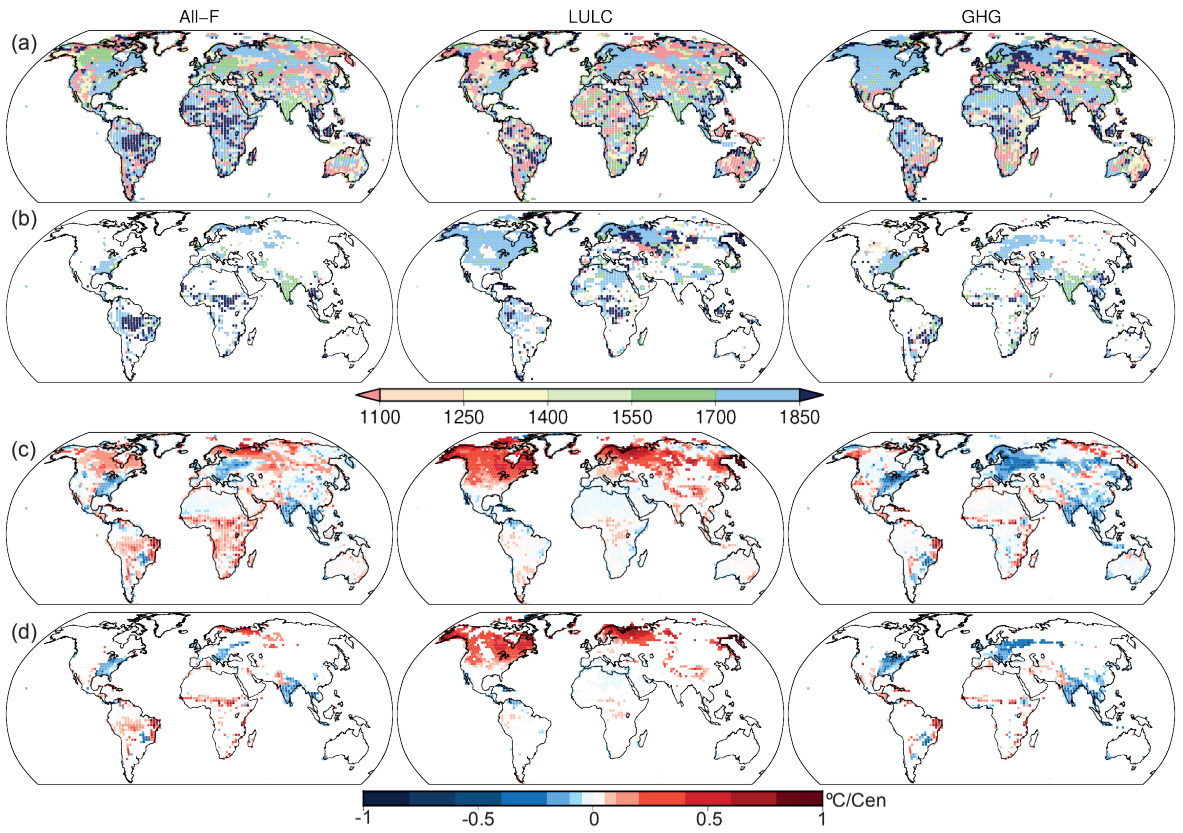

Figure 4.10: Spatial distribution of the two-phase regression results for SAT minus GST in the ALL-F_2, LULC$_1$ and GHG$_1$ ensembles as examples: (a) dates (years) of change; and (c) trends after the year of change. (b) and (d) are the same as (a) and (c) but only for areas where at least 80% of the ensemble members show significant changes. Trends before the year of change (not shown) are not significant in any of the ensembles. Significance (p<0.05) is obtained based on an F test (t test) for year of change (trends) following Solow (1987). Significance also accounts for autocorrelation (Trenberth, 1984). Note that the spatial window has been modified to enhance visualization of land areas.

Regionally, several circumstances account for impacting SAT–GST long-term coupling. On the one hand, decreasing SAT–GST differences over land may emerge from two conditions. Firstly, when there is a higher warming rate of SAT relative to GST, as depicted in Figure 4.9a,b over the northernmost part of North America, Fennoscandia, northeast Russia and some areas of central Eurasia. Secondly, when there is a cooling of both SAT and GST but the latter decreases at a higher pace as described in Figure 4.9a,b for the northeastern Brazilian region and some areas of Africa. These two scenarios are represented in Figure 4.10c for the ALL-F, with positive trends over these regions after the change. On the other hand, the increase in the SAT–GST difference either arises from the effect of rising GST in the presence of stable/decreasing SAT or due to the higher warming rate of GST relative to SAT. The former case is displayed in Figure 4.9a,b for central and eastern European areas and the eastern US, whereas the latter is found over the Indian subcontinent and southeastern Asia. Note that both cases are represented in Figure 4.10c (ALL-F) with negative trends.

Trends from the GHG-only and the LULC-only ensembles help understand the relative contributions to the long-term variations seen in the ALL-F simulations. For instance, the GHG-only ensemble shows similar positive trends to the ALL-F (Figure 4.10c) over northern North America, Fennoscandia, northeast Russia and central Eurasia with a much larger magnitude and geographical extension. Correspondingly, negative trends after the change in the LULC-only ensemble are comparable to those in the ALL-F for central and eastern Europe, the eastern US, the Indian subcontinent and southeastern Asia. Additionally, the positive values over Brazil and central and southern Africa in the ALL-F are also depicted in the LULC-only ensemble. Most of these changes are robust in 80% of the ensemble members (Figure 4.10d).

Interestingly, the two-phase regression analysis does not expose any variation in the SAT–GST long-term relationship over southeastern China, where the linear trends during the industrial period show a relatively strong decrease in both SAT and GST in the ALL-F and the OZ/AER-only ensembles (Figure 4.9a,b,g,h). Furthermore, the linear trend of SAT–GST differences during industrial times for the OZ/AER-only simulations (Figure 4.11) does not exhibit any SAT–GST decoupling over this region either. This suggests that the dominant effect of OZ/AER forcing on the SAT and GST responses over this region is not affecting their long-term coupling. Nonetheless, Figure 4.11 illustrates some interesting aspects of the SAT–GST relationship in the OZ/AER-only ensemble, such as the negative contribution to the SAT–GST trends over North America, northern Europe, the Tibetan Plateau and Central Asia. Additionally, the positive trends over northern Siberia are also notable, as well as the positive values over some relatively small areas of central and eastern Africa, the coast of Angola and

eastern Brazil. Although the bulk of these SAT–GST responses depicted in Figure 4.11 do not translate to SAT–GST long-term decoupling in the ALL-*F*ensemble, they play an important role in either counteracting the influence of other external forcings or contributing to decoupling-related processes over some regions.

The following paragraphs aim to provide insight into the relative contribution of the individual forcings and the associated physical mechanisms to the variations of the long-term SAT–GST association detected in Figures 4.10 and 4.11.

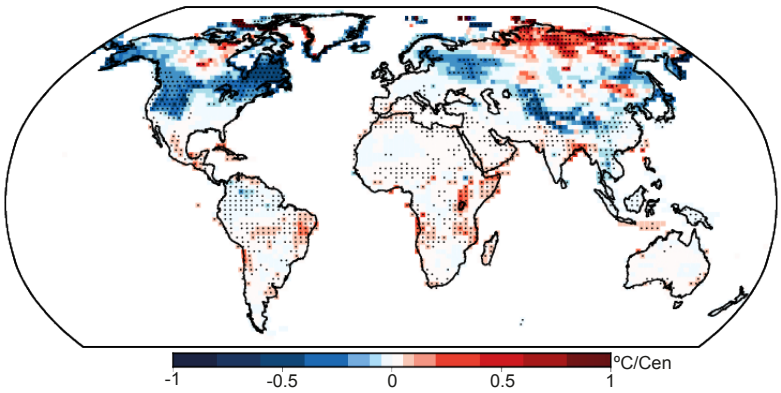

Figure 4.11: Spatial distribution of the linear trends for SAT minus GST in the OZ/AER-only ensemble. Trends are indicated in Celsius century^{-1}. Dots indicate agreement in both of the ensemble members

In the cases of the long-term variations due to the LULC influence, changes in vegetation cover alter the radiative fluxes and water cycling at the surface due to the modification of the physical properties such as albedo, roughness and evapotranspiration (Pongratz et al., 2010). Figure 4.12 gives an example of how long-term changes in the energy fluxes at the surface due to LULC changes impact the SAT–GST coupling at long timescales. It shows the 31-year low-pass filter outputs of SAT, GST, reflected shortwave radiation (RSW) and SHFLX evolution for a characteristic grid point over the Great Lakes region (US) where warming of GST relative to SAT is simulated in CESM-LME during the industrial period. Results are shown for one of the members of the ALL-*F*, LULC-only, GHG-only and OZ/AER-only ensembles. Around 1800 CE, the SAT tends to decrease whereas GST tends to increase in both the ALL-*F* and LULC-only simulations, while GHG-only and OZ/AER-only simulations do not display the

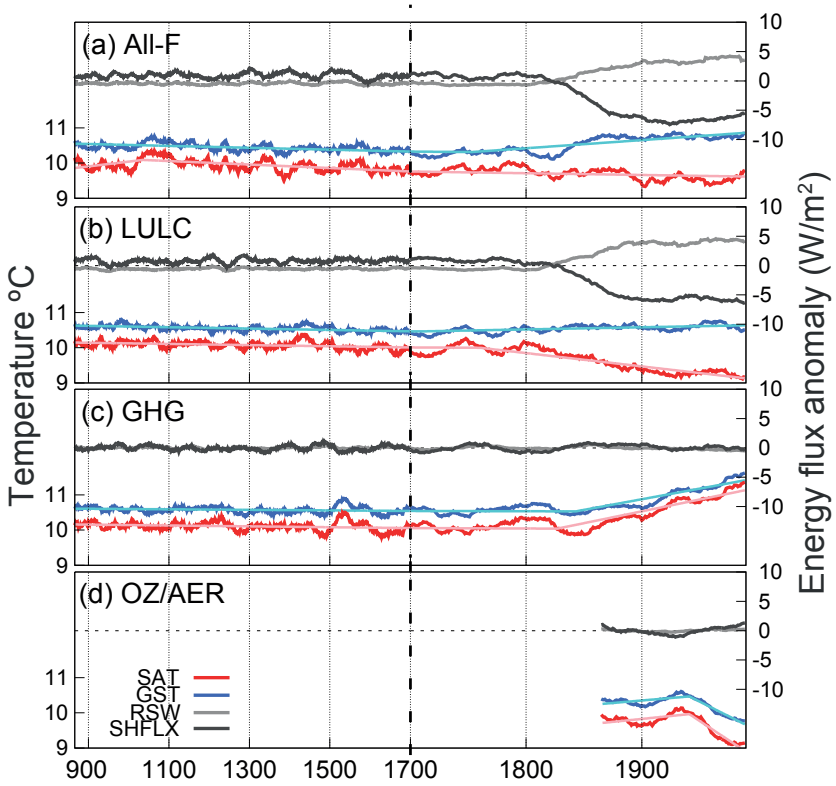

Figure 4.12: LM evolution of SAT, GST, reflected shortwave radiation (RSW) and surface sensible heat flux (SHFLX) for a grid-point located at 40°N,82°W at the south of the Great Lakes, U.S. in the ALL-F_2, LULC$_1$, GHG$_1$ and OZ/AER$_1$ simulations. The left axes correspond to SAT and GST, while the right axes to the energy fluxes at the surface. Note that for the energy fluxes the anomalies with respect to 850–2005 CE are shown, whereas for temperature absolute values are presented. All series are 31-yr moving average filter outputs. For SAT and GST the result of the two-phase regression model is displayed with thin solid lines. Note the change in timing of the x axis after 1700 CE.

same behavior that produces larger differences between SAT and GST represented by negative trends in Figure 4.10c. At the same time, RSW and SHFLX exhibit large long-term variations in the ALL-F and LULC-only simulations.

Therefore, this modification of the long-term SAT–GST relationship is clearly a response to LULC changes. Such variations in the surface energy fluxes over this region are likely response to vegetation replacement from forested areas to grassland or croplands. Forested landscapes dissipate SHFLX more efficiently to the atmosphere due to higher surface roughness than open fields (Jackson et al., 2008). Also, lower vegetation types have a higher reflectivity than forests. All of the previously listed factors contribute to SAT decreases over these regions, especially in DJF. Furthermore, deforestation at mid- and high-latitudes tends to positively feedback with increases in snow cover (Anderson et al., 2011). These types of changes in LULC contribute to increasing albedo, which is reinforced by changes in snow cover at these latitudes. Additionally, higher DJF snow cover tends to increase the insulation of the soil from the cold overlying air. The combination of these mechanisms leads to the observed temperature response of the SAT and GST. This particular LULC process is important for corrupting the SAT–GST coupling at timescales relevant for the borehole theory (centennial) since the thermal signature recorded in GST during the industrial period would not be representative of the past long-term SAT variations in regions where this effect is dominant.

For the areas of central and eastern Europe, where a GST warming relative to SAT is also observed in Figure 4.10c, the mechanisms are similar to those described in Figure 4.12, because these areas were also subject to an intense transformation from forested areas to cropland prior to the beginning of the industrial period according to the LULC forcings considered in the CESM-LME (Pongratz et al., 2008; Hurtt et al., 2009).

Changes in vegetation cover are also important for the long-term SAT and GST temperature differential response over tropical regions, although the driving mechanisms are different from those at mid- and high-latitudes (Lee et al., 2011) and they deserve to be considered. In the northeast of Brazil, both SAT and GST have negative trends in both the ALL-F and the LULC-only ensembles during the industrial period (Figure 4.9a,b,e,f). However, the decrease of GST is much larger than that of SAT as represented in Figure 4.10c with positive trends after the change in both ensembles.

Figure 4.13 shows the temporal evolution of SAT, GST, RSW and LHFLX for a grid point located at the northeast of Brazil. At the end of the 18th century, GST drops sharply whereas SAT slightly decreases. Likewise, RSW and LHFLX experience significant changes at the same time as a result of the modification of the surface characteristics. Such changes are present solely in the ALL-F and LULC-only simulations whereas GHG-only and OZ/AER-only ones show no differences in their evolution. The changes observed at this location in the energy fluxes likely correspond to transitions from open lands to a forested area (refor-

estation or afforestation) leading to lower albedo and higher evapotranspiration rates as is shown in Figure 4.13. This situation leads to an apparent long-term cooling of GST relative to SAT at this location. The temperature response over this area is influenced by different mechanisms. First, the conversion from lower- to higher-type vegetation reduces the solar radiation that impinges on the surface and GST decreases by a radiative effect. Second, forested lands usually have lower albedo thus absorb more shortwave radiation (Zhao and Jackson, 2014). This surplus of energy is balanced by the increase in transpiration, and consequently, GST decreases also by a non-radiative process. The latter is especially important in humid climates (von Randow et al., 2004) as the one in this example. According to this, there is a net heat loss at the ground surface with a higher decrease in GST relative to the overlying air. The SAT–GST coupling becomes strong again after some period when the new vegetation cover reaches a stable state by the mid 20th century. For the areas of Africa, where there is also a cooling of GST relative to SAT (Figure 4.10c in the ALL-F and LULC-only ensembles), the mechanisms are comparable to those described for Figure 4.13.

Over some of these tropical regions, there is also a contribution from the OZ/AER forcing to the SAT–GST response. Note that Figure 4.11 shows positive trends over Uganda, the coast of Angola and over eastern Brazil, which is also noticeable in Figure 4.10c. In these regions, the incoming solar radiation is reduced due to the effect of aerosols, which are an important element for cloud formation and contribute to a higher reflectivity of solar radiation (Tao et al., 2012). The description of the specific processes related to aerosol–cloud interaction goes beyond the scope of this study. Therefore, only the influence on the energy balance at the surface is addressed. As lower shortwave radiation impinges the surface, the energy gain decreases and the ground surface heating is consequently lower. The reduction in the energy gain at the surface is compensated for by a lower dissipation via sensible heat, whereas the fluxes of latent heat remain relatively constant or even increases in some areas due to higher moisture as a result of increased precipitation. This mechanism is illustrated in Figure 4.14 for a grid point located in the Uganda region. Note in the ALL-F simulation, the reduction in incoming shortwave radiation (SSW) after 1900 CE leads to a decrease in SHFLX. Interestingly, for this location, LHFLX experiences an increase simultaneously due to increased precipitation (not shown); thus, this provides a source of moisture for evapotranspiration. This situation leads to a higher GST decrease relative to SAT due to a net loss of energy at the ground surface. Similar SAT and GST responses are observed in the OZ/AER-only and LULC-only simulations, whereas in the GHG-only SAT and GST increase.

The SAT–GST decoupling processes described above for individual grid points are also important at larger spatial scales. Figure 4.15a shows an extension of

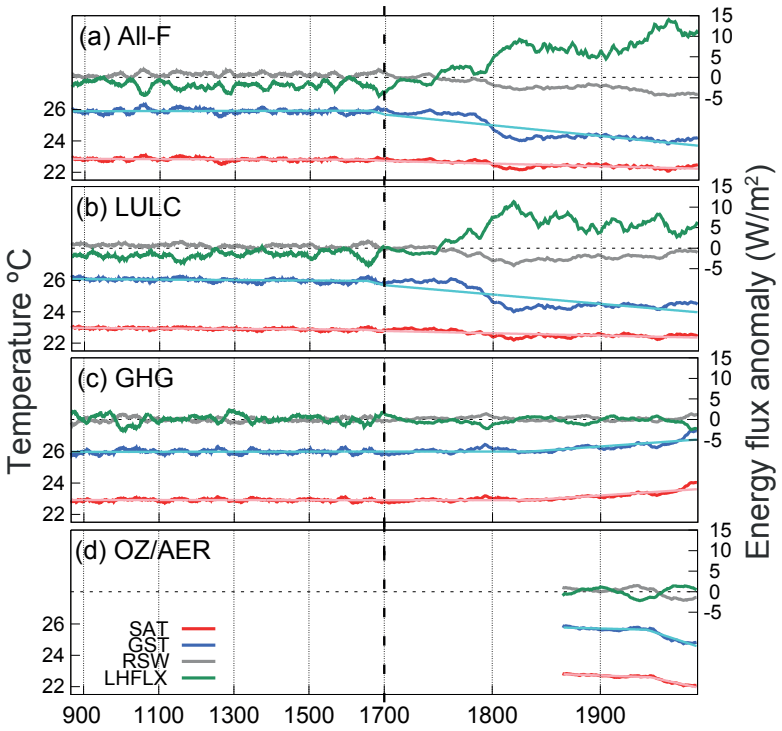

Figure 4.13: The same as in Figure 4.12 but for a grid point located at 12°S,40°W in the northeast of Brazil. LHFLX is represented instead of SHFLX.

the mechanism depicted in Figure 4.13 including a larger area over the northeast of Brazil (between 1-11° S and 47-35° W). The negative trend since the 18th century is less accentuated for SAT than for GST (-0.14 and -0.53 °C century^{-1} respectively in the ALL-*F* simulation and -0.07 and -0.33 in the LULC-only ensemble) indicating a strong contribution of past LULC changes. The same air–subsurface temperature response occurs in other tropical and subtropical areas such as the east of Africa.

The regional analysis is extended to southeastern China to illustrate additional information about the influence of different external forcings on the SAT–GST relationship. As previously discussed, the negative trends for both SAT and GST over this region during the industrial period represented in the ALL-*F* and OZ/AER-only ensembles (Figure 4.9a,b,g,h) do not entail a corruption of the

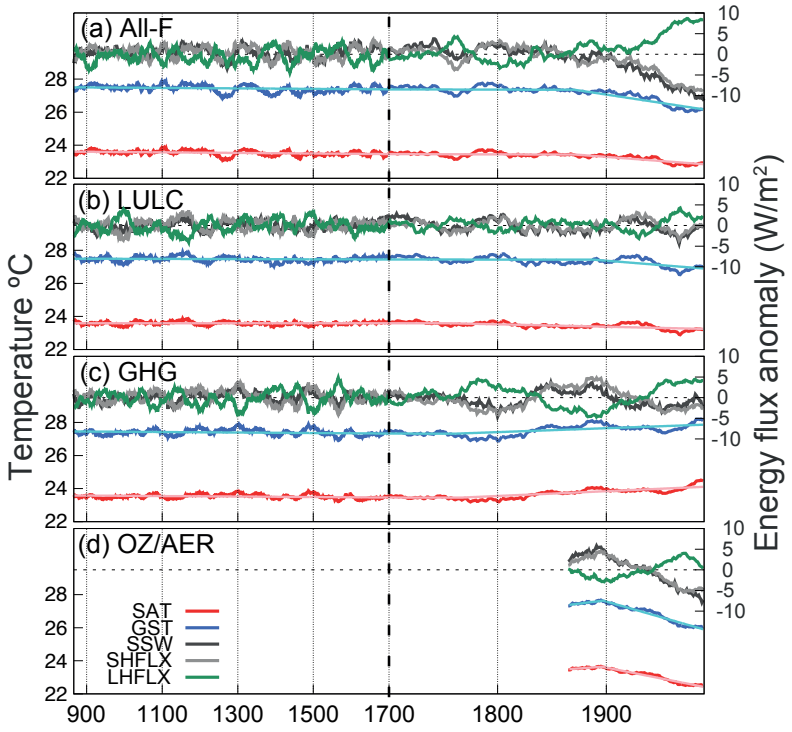

Figure 4.14: The same as in Figure 4.12 but for a grid-point located at 2.5°N,33°E over Uganda. Note that incoming short wave radiation at the surface (SSW) is represented instead of RSW and both LHFLX and SHFLX are shown.

SAT–GST long-term coupling. Figure 4.15b allows for insight into this particular effect and the role of the OZ/AER forcing on the air and soil temperature responses over this region during industrial times. The negative long-term trend within the industrial period is only seen in the ALL-*F* and the OZ/AER-only simulations. Similarly, in both of these simulations, there is a reduction in the RSW and in both the LHFLX and SHFLX. Conversely, the SAT and GST evolution during industrial times in the GHG-only and LULC-only simulations does not follow the same path depicted in the ALL-*F*, highlighting the dominant influence of OZ/AER forcing.

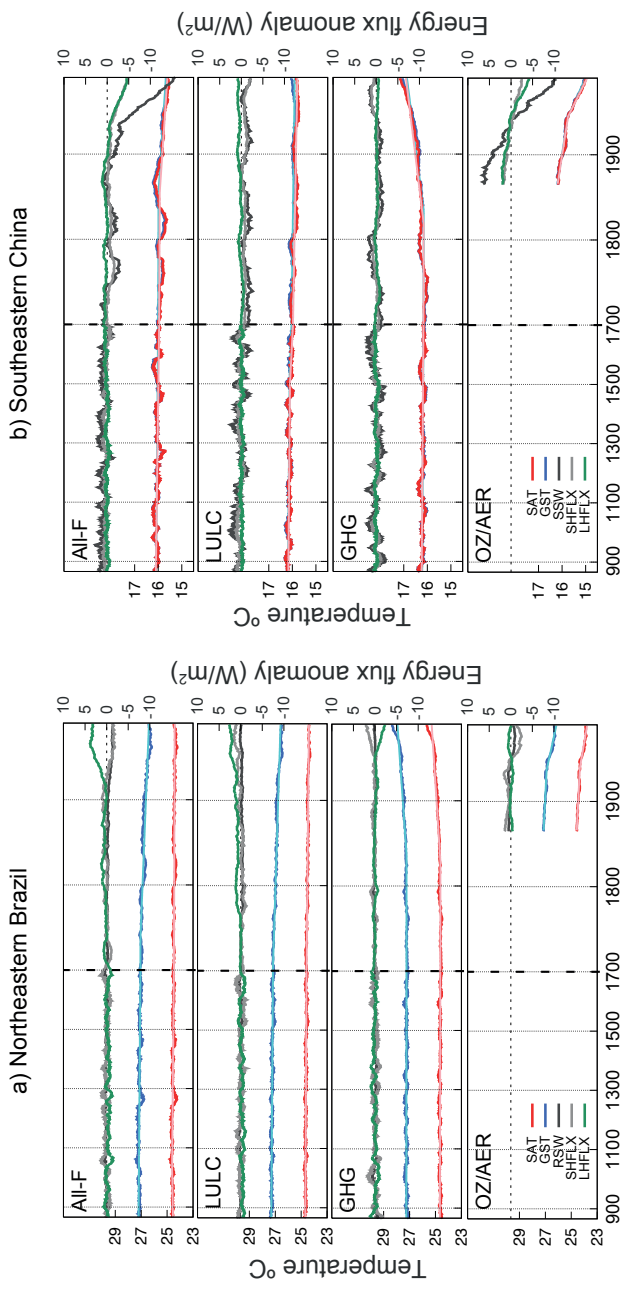

Figure 4.15: The same as in Figure 4.12 but for two different regions: northeast of Brazil between 1°S-11°S and 47°W-35°W (left) and southeast of China between 22°N-32°N and 103°E-122°E (right). For the area of northeast Brazil RSW is represented whereas for the area of southeast China SSW is shown instead. LHFLX and SHFLX are shown in both cases.

In southeastern China, the variations of the energy fluxes at the surface depend on the reduction of the incoming shortwave radiation as a response of the anthropogenic aerosol–cloud interaction rather than by modifications of the land surface properties. Therefore, the decrease in the energy that impinges the surface is balanced by decreasing both the sensible and latent heat fluxes. Hence, the air–soil interactions are not significantly altered and the SAT–GST relationship remains stable.

Although LULC changes induce the majority of the important long-term variations in the SAT–GST relationship at regional and local scales observed in the ALL-F ensemble (Figure 4.10c), there are some regions in which the GHG forcing is the main driver for long-term SAT–GST decoupling. For instance, the positive trends after the change over Fennoscandia, northeast Russia and the north of North America observed in the ALL-F ensemble can be explained to a large extent by the influence of GHG-only ensemble inasmuch as a broadly similar picture is portrayed over these regions in both ensembles (Figure 4.10c).

In the GHG-only ensemble case, the strong warming of the SAT relative to GST over these regions is driven by the increasing air temperature during industrial times due to the positive radiative forcing of GHGs in the presence of a considerable long-term reduction in simulated snow cover. González-Rouco et al. (2009) showed that such a scenario would lead to a higher exposure of soil to cold winter air; therefore, the soil would partially record colder temperatures, which had previously been prevented by the snow cover insulating effect. In the ALL-F ensemble, this effect is damped as additional forcings that keep the snow cover relatively constant during industrial times are considered. For instance, the OZ/AER contribution forcing is particularly important for counteracting the effect of GHGs as it leads to colder climate conditions due to its negative radiative forcing. Note the strong negative trends in Figure 4.11 over North America, northern Europe and the Tibetan Plateau that partially balance the effect of the GHGs over these regions. Nonetheless, there is still an overall SAT warming relative to GST in the ALL-F since the relatively stable snow cover insulates the soil from a warmer SAT, reducing the overall offset between them (Bartlett et al., 2005). Additionally, other processes can play some local roles in SAT–GST changes, such as CO_2 generating increases in leaf area index (e.g., Amazon rainforest), which leads to higher evapotranspiration (Mankin et al., 2018) and GST cooling relative to SAT.

Figure 4.16 gives further insights into the interactions between anthropogenic forcings, their influence on global snow cover and consequently on the long-term SAT–GST relationship in the CESM-LME. The LM evolution of global SAT minus GST (top) and the annual global snow cover (bottom) for the ALL-F, LULC-only, GHG-only and OZ/AER-only simulations are shown for one member

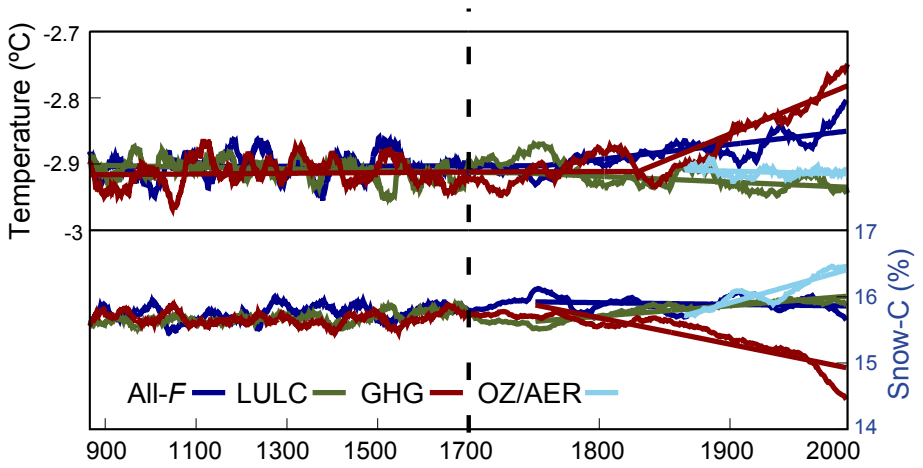

Figure 4.16: Top. LM evolution of the global SAT–GST offset for the ALL-F_2, $LULC_1$, GHG_1 simulations. Straight lines indicate the long-term trend during the LM from a two-phase regression analysis except for the OZ/AER_1 results that indicate a linear trend in the industrial period. Bottom. LM evolution of global snow cover percentage for the same ensemble members. Straight lines indicate the long-term trend within the 1750–2005 period except in the case of OZ/AER that only covers the 1850–2005 period. The period from 1750–2005 was selected in order to match the time span when the SAT–GST offset experiences the variation in the ALL-F_2. All series are 31-year moving average filter outputs. Note the change in timing on the x axis after 1700 CE.

of each ensemble. Similar results are obtained if other members are selected. On the one hand, when the GHG-only ensemble is considered, the global SAT–GST offset experiences a sharp long-term decrease in absolute value at the start of the industrial period and a strong long-term reduction in global snow cover. In fact, the correlation between changes in the SAT–GST offset and snow cover in the GHG-only ensemble member is high -0.93 (p<0,05). On the other hand, the overall effect of the LULC-only ensemble is a relatively small increase in the global snow cover mainly due to deforestation at mid and high latitudes as well as due to the negative radiative forcing of LULC as Earth's albedo increases (Myhre et al., 2013); this leads to a small increase in the global SAT–GST offset. Similarly, the OZ/AER-only ensemble shows an increase in global snow cover while the SAT–GST offset in industrial times exhibits a relatively slight increase. The interaction

between different external forcings in the ALL-F ensemble leads to a relatively stable snow cover during the industrial period since the sharp decrease in snow, induced by the GHG forcing, is partially compensated for by the counteracting effect of the LULC and OZ/AER forcings. Additional forcings such as volcanic eruptions may also contribute to counteracting GHG effects at multi-decadal timescales (not shown). Consequently, in the presence of a warmer climate, there is a difference in the warming rate of SAT and GST in industrial times at a global scale in the ALL-F ensemble member (0.25 and 0.18 °C century^{-1}, respectively). This scenario leads to the net effect of a long-term decrease in the SAT–GST differences starting around 1800 CE as discussed in Figure 4.2 and as is also evident in Figure 4.16.

4.3 Conclusions

In this chapter, the stationarity of the coupling between SAT and GST temperatures as simulated by the CESM in an ensemble of experiments spanning the LM has been evaluated. The analysis is motivated by previous literature (González-Rouco et al., 2006, 2003, 2009; García-García et al., 2016) that addresses the realism of the borehole hypothesis for climate reconstruction, namely, that SAT and GST varies synchronously and that reconstructing past GST changes from borehole temperature profiles is a good proxy for past SAT variations. The use of the CESM-LME allows for analyzing the influence of forcing changes on the SAT–GST covariability, both individually and as a group, by considering the different all-forcing and single-forcing ensembles. Additionally, having several experiment ensemble members for each given forcing type allows for disentangling the effects of internal variability from the forced response. Ultimately, the coupling between SAT and GST is assessed globally and at regional/local scales. This assessment requires the consideration of different mechanisms that contribute to SAT–GST variability within different climate types. In doing this, a variety of factors and conditions that contribute to the surface energy balance via different mechanisms are provided.

The CESM-LME shows that at global scale the SAT–GST coupling is strong above multi-decadal timescales since GST tracks SAT throughout the LM, as found in previous studies. However, despite the strong coupling, the CESM-LME also reflects that the SAT–GST relationship has not remained constant throughout the whole LM at these spatiotemporal scales. Hence, the nature of such variation is evaluated.

Globally, snow cover is the most important agent in modulating the connection between SAT and GST. Therefore the variation of the SAT–GST relationship described by the CESM-LME simulations at a global scale should be, in principle, driven by variations of global snow cover. Nevertheless, the simulated snow cover remains relatively stable when SAT–GST coupling varies; thus, this change cannot be solely explained by the influence of the snow cover. With this in mind, different processes that may influence the SAT–GST relationship at different spatiotemporal scales were explored in detail. Firstly, the processes acting at seasonal timescales identified from a spatial analysis of the SAT–GST differences and correlations have been addressed. Secondly, the long-term evolution of the SAT–GST relationship is evaluated in the ALL-*F*, LULC-only, GHG-only and OZ/AER-only ensembles.

Several processes over different regions relevant during either DJF or JJA play an important role in impacting the SAT–GST coupling, such as snow cover over mid- and high-latitudes, discontinuous snow cover over the Tibetan Plateau region and seasonal variations in the energy fluxes at the surface. Although these processes are important for disrupting the SAT–GST relationship at seasonal scales, they have no implications on the long-term coupling if they are stationary. Nonetheless, if they experience variations with time the SAT–GST long-term relation may be impacted.

As discussed in Section 4.2.1, some anthropogenic external forcings can impose long-term variations on processes that regulate the relationship between SAT and GST. Among them, the LULC changes are the most important ones since they modify the energy fluxes at the ground–air interface, and consequently corrupt the SAT–GST coupling locally and regionally at various timescales. One example is the response to the deforestation processes triggered by the expansion of agriculture mainly during the industrial period at mid- and high-latitudes, where SAT and GST long-term coupling is impacted due to the variations in the albedo, surface roughness and hydrology. Similar decoupling processes related to LULC changes are found over different regions around the globe, such as those described in Section 4.2.1 over northeastern Brazil, some areas over Africa and the Indian subcontinent. All of them driven by the long-term modifications of the energy fluxes at the surface either from increased evapotranspiration, reduced energy dissipation via sensible heat and others factors. Besides these kinds of decoupling processes induced by individual forcings, the interactions of various mechanisms and feedback from different external forcings can also influence the long-term SAT–GST relationship at different spatial scales. For instance, the effect of GHGs leads to a reduction of the snow cover during industrial times that is counterbalanced by the opposite effect of both LULC and OZ/AER forcings. Consequently, the snow cover remains relatively stable over some regions during

the industrial period in the presence of a warmer climate. This situation leads to a difference in the SAT and GST long-term evolution during the industrial period, since the snow cover insulates the soil from a warmer SAT. This effect is present over the NH high-latitudes of North America, Fennoscandia and northeastern Russia. Indeed, on a global scale, the combination of a steady snow cover under warmer climate conditions is the dominant effect for explaining the variations in the long-term SAT–GST relationship.

If different ESMs/GCMs would be considered the analysis may yield different levels of impact. Overall, some consistency in the impacts among the available simulations of the LM with different GCMs may be expected, since the external forcings considered in the PMIP3/CMIP5 LM simulations are similar (Schmidt et al., 2011, 2012) and should have a similar contribution to the extent that the impacts on the SAT–GST long-term relationship is mainly driven by the external forcings. Nevertheless, not all model simulations consider exactly the same set of forcings. Some consider land use land cover changes and aerosols, and other do not (Masson-Delmotte et al., 2013). In addition, different model climate sensitivities and differences in the representation of surface processes (e.g., snow cover or soil moisture) may also contribute to produce different responses. Even if the role of external forcings should be dominant on a long-term context, addressing this issues from a multi-model ensemble approach would help understanding the uncertainties associated to all the factors of variations described above. The CMIP6/PMIP4 (Eyring et al., 2016; Jungclaus et al., 2017) offers an opportunity for exploring better sensitivity to different models since largest ensembles of the LM with a more systematic sampling of forcings and processes would be expected.

The findings described in this chapter indicate that the assumption of a strong relationship between SAT and GST may be impacted from local to regional scales by different mechanisms, especially by the influence of LULC changes due to the modification of the energy balance at the surface. Therefore, the interpretation of temperature reconstructions from borehole measurements at these spatial scales must consider LULC changes as a source of possible bias. The effects of additional external forcings may also exert some influence on processes such as variations in the snow cover, hydrology and other land surface properties, that may in turn feedback on the SAT–GST long-term coupling. At global scale, the influence of such local and regional decoupling processes is only ca. 0.05 °C. Hence the SAT–GST coupling at this spatial scale is supported by the CESM-LME.

5

Methodological and physical biases in borehole temperature reconstructions*

Chapter 4 provides the main motivation to advance in the investigation of the full set of potential sources of bias in borehole-based surface air temperature reconstructions since they show that long-term SAT–GST decoupling may exist mainly due to the influence of changes and associated feedbacks in GHGs and LULC external forcings. Accordingly, the inversion of BTPs in which the SAT–GST long-term coupling is corrupted would produce unreliable estimations of the past SAT variations. Hence, this chapter provides a follow up on Melo-Aguilar et al. (2018) findings. The analyses presented in this chapter expand previous assessments and present a set of PPEs in which the borehole-based reconstruction strategy is implemented using LM simulations from the CESM–LME as in Chapter 4. Previous analyses regarding the methodological constraints that can lead to bias estimations of the past GST evolution are updated and extended to new issues. This includes using in the experimental design the actual spatial distribution of BTPs and dates and depths of the borehole logs (Huang and Pollack, 1998). Likewise, the contribution from the physical mechanisms that corrupt the long-term SAT–GST coupling assumption to a biased representation of past SAT variation from GST reconstructions is systematically addressed. Finally, the combined effect of both methodological and physical sources of bias is assessed, considering a range of global to regional scales.

Section 5.1 provides a description of the strategy implemented to reproduce the actual borehole-based temperature reconstructions. The assessment of the methodological issues influencing the estimates of past GST at large spatial

* The main contents of the chapter are included in the published research article:
- Melo-Aguilar, Camilo, Jesús Fidel González-Rouco, Elena García-Bustamante, Norman Steinert, Johann H. Jungclaus, Jorge Navarro and Pedro J. Roldn-Gmez 2020: Methodological and physical biases in global to subcontinental borehole temperature reconstructions: an assessment from a pseudo–proxy perspective *Climate of the Past*, **16**,453–474. 10.5194/cp-16-453-2020.

scales is developed in Section 5.2.1. This includes a thorough evaluation of differences between the estimation of trends from the different setup of pseudo-reconstructions and the simulated GST. Further, the specific effects of each of the methodological constraints (i.e., uneven spatial distribution, different timing and depths of the borehole logs) at global to sub-continental spatial scales are explored. For the assessment of methodological issues, only the ALL-F simulations are considered since the objective is testing the overall performance of the technique in retrieving the GST signal in a realistic context. Section 5.2.2 deals with the contribution of long-term SAT–GST decoupling to the reliability of GST reconstructions from BTPs in representing the past SAT variations. In this part, in addition to the ALL-F set of simulations, the GHG- and LULC-only experiments are also included as it has been shown that these external forcings are the main drivers to produce long-term SAT–GST decoupling (Chapter 4). As in the case of the methodological issues, the analyses developed in this part are constructed at global to sub-continental scales.

5.1 Experimental design

The theoretical basis for the borehole temperature reconstruction states that the subsurface contains a thermal signature of the past surface temperature variations due to the superposition of the downward temperature signal propagating onto the background geothermal gradient. Therefore, the inversion of BTPs can yield information of the past surface temperature changes. The information of the last 500 to 1000 years is retained within the upper few hundred meters of the subsurface (Beltrami and Bourlon, 2004). Thus, BTPs of at least 300 m depth are required to retrieve the past 500 years (Jaume-Santero et al., 2016), whereas deeper profiles of at least 500 m are necessary to retrieve information of the complete LM (Pollack and Huang, 2000).

Within the model world, the depths of BTPs are limited to the land model depth. Although the CLM4 has the deepest BBCP among the current land surface models, it is still too shallow to directly provide such deep BTPs to account for LM temperature variations (see Table 2.4). Consequently, these profiles must be synthetically generated from the surrogate reality of the simulated world as in González-Rouco et al. (2006). The following section provides an overall description of the models employed in this study to simulate the synthetic BTPs as well as for the inversion of the resulting profiles. Further details can be found in Mareschal and Beltrami (1992).

5.1.1 Forward and inverse models

The BTPs are determined by the combination of the geothermal heat flux, a reference ground temperature and the temperature perturbation $T_t(z)$ induced by the surface temperature variations, as follows:

$$T(z) = T_0 + q_0 R(z) + T_t(z), \tag{5.1}$$

where q_0 represents the surface heat flow density, $R(z)$ is the thermal depth and T_0 is a reference ground temperature. In the forward model, the quasi-steady-state component, $T_0 + q_0 R(z)t$, can be set equal to 0 because the aim is to derive $T(z)$ as a function of only the past surface temperature variations. The forward model thus determines the transient perturbation component $T_t(z)$, which can be thought of as the anomaly with respect to the quasi-steady-state thermal regime.

The propagation of the surface temperature signal within the subsurface is controlled by the one-dimensional time-dependent heat conduction equation (Carslaw and Jaeger, 1959):

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2}, \tag{5.2}$$

with κ as the thermal diffusivity and z and T as depth and temperature, respectively. Equation (5.2) is solved for an instantaneous temperature change at time t before present as:

$$T_t(z) = \Delta T erfc\left(\frac{z}{2\sqrt{\kappa t}}\right), \tag{5.3}$$

where $erfc$ is the complementary error function. $T_t(z)$ can be modeled by considering the temperature variations at the surface as a series of K time step temperature changes. In this way, each step imprints a thermal signature in the subsurface that is merged to the signature of the previous step. Thus, $T_t(z)$ is given by (Mareschal and Beltrami, 1992):

$$T_t(z) = \sum_{k=1}^{K} \Delta T_k \left[erfc\left(\frac{z}{2\sqrt{\kappa t_k}}\right) - erfc\left(\frac{z}{2\sqrt{\kappa t_{k-1}}}\right)\right], \tag{5.4}$$

where ΔT_k are the surface temperature changes for K time intervals, each value representing an average over time $(t_k - t_{k-1})$.

Equation (5.4) is used to create synthetic BTPs, using LM surface temperature annual anomalies from the CESM-LME as the upper boundary condition. Although the synthetic temperature profile represents only the transient perturbation component, $T_t(z)$, of the BTP, it will be denoted as BTP thorough the document to avoid confusion. $T_t(z)$ is evaluated at every 1 m depth interval up to a depth of 600 m in order to accommodate for the propagation of the LM surface temperature variations. Subsequently, the upper 20 m of the resulting BTPs is removed in order to avoid the influence of the annual signal and reproduce realistic depths of the water table (Jaume-Santero et al., 2016). The thermal diffusivity used in the geothermal models is 1.5×10^{-6} m^2 s^{-1}, obtained from the values of the bedrock thermal conductivity (3.0 W m^{-1} K^{-1}) and volumetric heat capacity (2×10^6J m^{-3} K^{-1}) of the CLM4 (Lawrence et al., 2011).

For the inversion of the synthetic BTPs, singular value decomposition (SVD; Mareschal and Beltrami, 1992) is applied to retrieve the past long-term surface temperature variations. For the present work, the inverse model consists of a series of 15-year step changes in surface temperature histories following García-García et al. (2016). Different parameterizations were also tested (e.g., 20-year or 25-year step changes; not shown) and showed results consistent with the 15-year discretization. The latter was finally selected because it yielded a convenient representation of the GST histories. Likewise, the number of retained singular values is also important. The presence of small singular values leads to an unstable solution dominated by noise, while the use of only a few principal components (PCs) result in a smoothed low-resolution solution (Beltrami and Bourlon, 2004). The selection of the number of singular values is done by setting an eigenvalue cutoff, from which smaller values are eliminated. In this study we have used a cutoff value of 1.5×10^{-1} from which the solution is the linear combination of the three leading modes. We have additionally explored the effect on the solution of retaining four and five PCs.

5.1.2 Pseudo-proxy set up

The assessment of the impacts of the methodological and physical issues described above on borehole temperature reconstructions, is done by designing two pseudo-proxy configurations. First, a so-called ideal borehole scenario (IBS) is considered, in which a BTP is simulated at every land model grid point up to a depth of 600 m. This scenario is also characterized by a homogeneous logging date at the end of the CESM-LME simulation period (2005 CE) at every grid point. The latter is achieved by driving the forward model with the annual temperature

anomalies calculated with respect to the 850–2005 CE mean. Subsequently, each of these BTPs is inverted, and a latitude weighted average is used to obtain the global mean. A similar approach has been used in previous works showing that under such an idealized configuration, the borehole technique is able to retrieve the simulated surface temperature variations at the global scale (González-Rouco et al., 2006; García-García et al., 2016). Hence, this scenario provides a benchmark experiment that will allow for an evaluation of the impacts of having some methodological constraints that mimic those in reality. The IBS scenario embraces two cases. On the one hand, the model soil temperature (ST) at layer 12 (ST_{L12} ∼7.8 m depth) is used as the upper boundary condition (IBS_{L12}). This case represents the ideal scenario in which to generate the pseudo-reconstructed GSTs. On the other hand, the 2m air temperature model output is employed (IBS_{SAT}) in order to obtain an ideal case of pseudo-reconstructed SATs. We use the ST_{L12} as the reference ST to force the IBS_{L12} forward model because in the upper 10 layers, the CLM4 is hydrologically active and it is interesting to keep this realism in the synthetic BTPs. Additionally, this depth is consistent with using annual resolution in the model data as at this level the annual wave is very damped in amplitude, and layers below start to filter out decadal timescales. Nevertheless, this decision does not influence results significantly.

Second, a more realistic arrangement of the available global borehole network is implemented (B_{mask}), including the actual spatial distribution of the borehole sites as well as their real log depths and dates. This information is obtained from the "Global Database of Borehole Temperatures and Climate Reconstructions" (Huang and Pollack, 1998). Only the BTPs deeper than 200 m are considered in this step since the focus is on multi-decadal to centennial timescales within the LM as in previous reconstructions works (e.g., Huang et al., 2000; Beltrami and Bourlon, 2004). After this selection, a total of 970 logs are retained. For this setup, each of the borehole sites is placed on a specific land model grid point according to their geographic information (Figure 5.1). If more than one borehole corresponds to the same model grid point, only one of them is retained, prioritizing the most recent one to allow for the climatic signal of the last few decades of the 20th century to be represented. Additionally, for the bulk of the cases, the most recent borehole records coincide with the deepest ones. We finally kept 301 grid points that are the nearest co-located with any of the real borehole locations. Once the spatial distribution of the borehole network is represented in the CESM-LME grid, the LM ST_{L12} series at each of these grid points is trimmed at the actual logging date according to the date distribution (Figure 5.1a). Then, the temperature anomalies are calculated with respect to the trimmed period mean. The resulting anomaly time series are subsequently used to force the forward model at each grid point, generating a pseudo-BTP that is shortened to the

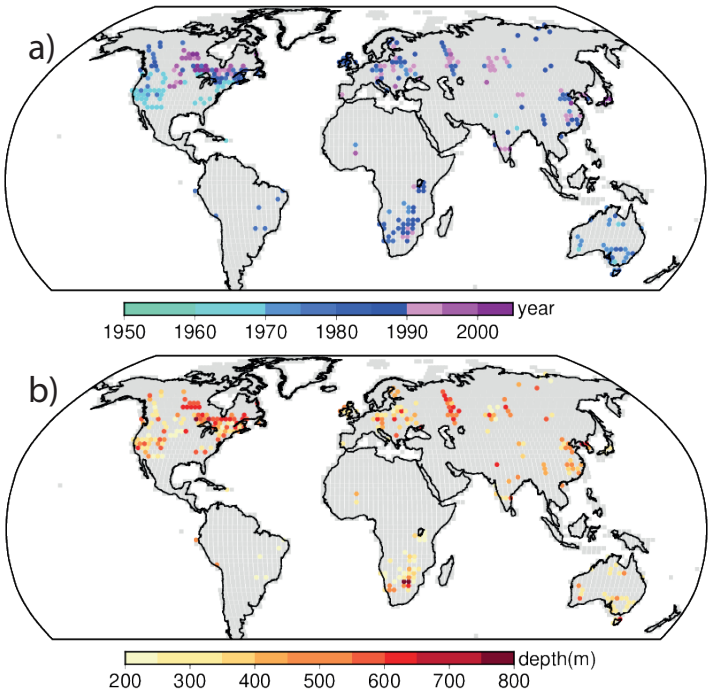

Figure 5.1: Spatial distribution of a) the logging date and b) the depths in the actual borehole network.

actual borehole depth (Figure 5.1b) in order to keep this configuration as realistic as possible. In this configuration, the inversion of the individual profiles yields information until the date when they were logged; therefore the reconstructed period is highly variable. The latter seemingly has an impact on the estimation of global averages since closer to present day the number of available sites decreases. Indeed ca. half of the borehole profiles were logged before 1980 CE, and only ∼20% of the sites have been measured after 1990 CE. In order to account for such variability in the logging dates, global and regional averaging is done on a yearly basis (Jaume-Santero et al., 2016).

The results from the idealized IBS and the more realistic B_{mask} configuration are then compared in order to evaluate the methodological limitations as well as the potential physical bias. As the methodological aspects are related to the

suitability of the spatiotemporal distribution of the borehole network to retrieve the past GST variations, IBS_{L12} allows for an assessment of these type of limitations when compared to B_{mask}; the differences between them informing about the effect of the methodological variants. This assessment is initially developed at the global scale, and then, it is extended to smaller spatial domains in order to address the implications on areas with different spatiotemporal borehole distribution. In this part of the analysis only the ALL-F ensemble is used. In this work, GST, which is ultimately the target signal of IBS_{L12}, is defined as the ST at the first soil layer (ST_{L1}; 0.007 m depth), following the same convention as in Melo-Aguilar et al. (2018).

The physical-related aspects and their influence on the estimates are subsequently explored via the comparison with IBS_{SAT}, because any deviation from the pseudo-reconstructed SAT would be the result of the physical SAT–GST decoupling. Moreover, while the differences between the IBS_{SAT}, and the IBS_{L12} cases would provide information about the individual contribution of the physical aspects in the interpretation of the SAT variations from the pseudo-reconstructed GST and the differences between IBS_{L12} and B_{mask} include the effect of the methodological sampling issues described above, the differences between IBS_{SAT} and B_{mask} allow for the estimation of the combined effects of both the methodological constraints and the physical processes. Besides the ALL-F ensemble, in this analysis the GHG- and LULC-only ensembles are also considered in order to estimate their specific contribution to the potential physical biases. Therefore, the previous PPE setups (IBS_{SAT}, IBS_{L12}, B_{mask}) are run for each of the simulations in the ALL-F and GHG- and LULC-only ensembles.

In order to allow for an easy identification of the different abbreviations and acronyms referenced along this chapter and their precise meaning, Table 5.1 contains a detailed description of them.

5.2 Results

The results of generating a variety of PPEs are reported herein. First, those related to methodological sampling issues (Section 5.2.1) and second, addressing SAT-GST coupling (Section 5.2.2).

Table 5.1: Abbreviations and acronyms used in this chapter.

Acronym	Meaning
GST	Simulated ground surface temperature defined as the first soil layer temperature (ST_{L1}; 0.007 m depth).
SAT	Simulated 2m air temperature. Original model output TREFHT.
IBS_{L12}	Ideal borehole scenario created from ST_{L12} as the boundary condition for the forward model.
IBS_{SAT}	Ideal borehole scenario created from SAT as the boundary condition for the forward model.
GST_{mask}	GST masked with the realistic representation of the variability of the spatio-temporal distribution of the global borehole network. In some cases GST_{mask} refers to the full spatial+date sampling whereas in other cases it refers only to spatial sampling (i.e. Fig 5.4).
B_{mask}	Realistic scenario of the borehole temperature inversions including sampling in space, time and depth. It may also refer to the cases in which the sampling is only in space or space+depth (i.e., Figs 5.2 and 5.4).

5.2.1 Methodological issues

The comparison of the simulated global GST anomalies within the IBS_{L12} and the B_{mask} pseudo-reconstructions in the ALL-F_2 and ALL-F_5 members of the ALL-F ensemble are represented as an example in Figure. 5.2a. These members were selected because they represent two possible results of the pseudo-reconstructed GST in the B_{mask} configuration, which are discussed herein. Similar results are obtained if other members are selected. In general, the IBS_{L12} pseudo-reconstruction reasonably reproduces the gross features of the low-frequency GST variations over the LM in the CESM-LME. For instance, a small non-significant multicentennial cooling from the MCA to the LIA can be detected, and the warming over industrial times is successfully captured in both cases. Within the LIA and back to the MCA, the filtering effect produced by the heat diffusion averages over intervals of multi-decadal and centennial warming and cooling. Changes in the past, like the simulated MCA warming, get damped in BTPs because of this effect, and a very smooth version of them is recovered by the reconstruction. Nevertheless, in model experiments that simulate larger MCA–LIA changes, the borehole reconstruction is able to recover somewhat warmer temperatures during

the MCA if the depth of the anomalies and the number of eigenvalues retained are adequate (González-Rouco et al., 2006, 2009).

Therefore, the IBS_{L12} qualitatively reproduces the long-term trends in both examples of Figure 5.2a. However, the masked inversion (B_{mask}) reproduces identical results in the case of ALL-F_5 but diverges from ALL-F_2 in the last few decades of the 20th century. This suggests that the results of B_{mask} can be simulation dependent, i.e., dependent on the different initial conditions and therefore on internal variability. To evaluate this, a more quantitative estimation of trends is needed, one that allows for a comparison of the idealized and masked pseudo-reconstructions and also a comparison of both of them with the simulated time series. To facilitate this, two approaches are proposed. One of them is to make a simple linear fit of the temperatures, either simulated or pseudo-reconstructed over a reference period. As this approach can potentially be affected by the different discretization of the simulated (annual resolution) and pseudo-reconstructed (15-year time steps) GSTs, a second strategy is applied in which GST series are transformed to 15-year averages that coincide with the time steps of the pseudo-reconstructed inversions over the time interval considered. Trends are then calculated by subtracting the mean values of the last and first 15-year steps and dividing the result by the total length of the reference time interval. This allows for a verification of the robustness of the results. We focus on the 20th century to evaluate trends because of the following: the bulk of the warming takes place in this period; results are less sensitive to the selection of the 15-year interval than if the 19th century is considered, due to the influence of natural (internal and forced) variability; and it offers the possibility to compare results with instrumental trends (Hartmann et al., 2013). Thus, Figure 5.2b shows results for linear fit to the 1900–2005 CE period and for trends calculated from the rates of change between 1890–1905 and 1990–2005 CE (15-year-diff in Figure 5.2b). Figure 5.2b shows the frequency distribution of trends calculated from both approaches for all members within the ALL-F ensemble. Box-and-whisker plots are shown for all possible scenarios considered herein, which include the following: GST, IBS_{L12}, GST masked with the realistic borehole configuration in space and time (GST_{mask}), B_{mask} as well as the differences among them (IBS_{L12}–GST, B_{mask}–GST_{mask}, GST–GST_{mask} and IBS_{L12}–B_{mask}). In the case of the B_{mask} series, the trend estimation considers the time interval from 1900 to the last available date.

Interestingly, the frequency distribution of trends within the ALL-F ensemble is similar for both strategies (15-year-diff and linear fit). The estimated global trends (GST in Figure 5.2b) show statistically significant values ($p < 0.05$) that range between 0.3 and 0.6 °C century^{-1} across the 13 simulations. Thus, internal variability has an impact on these trends estimates. The significance of the trends

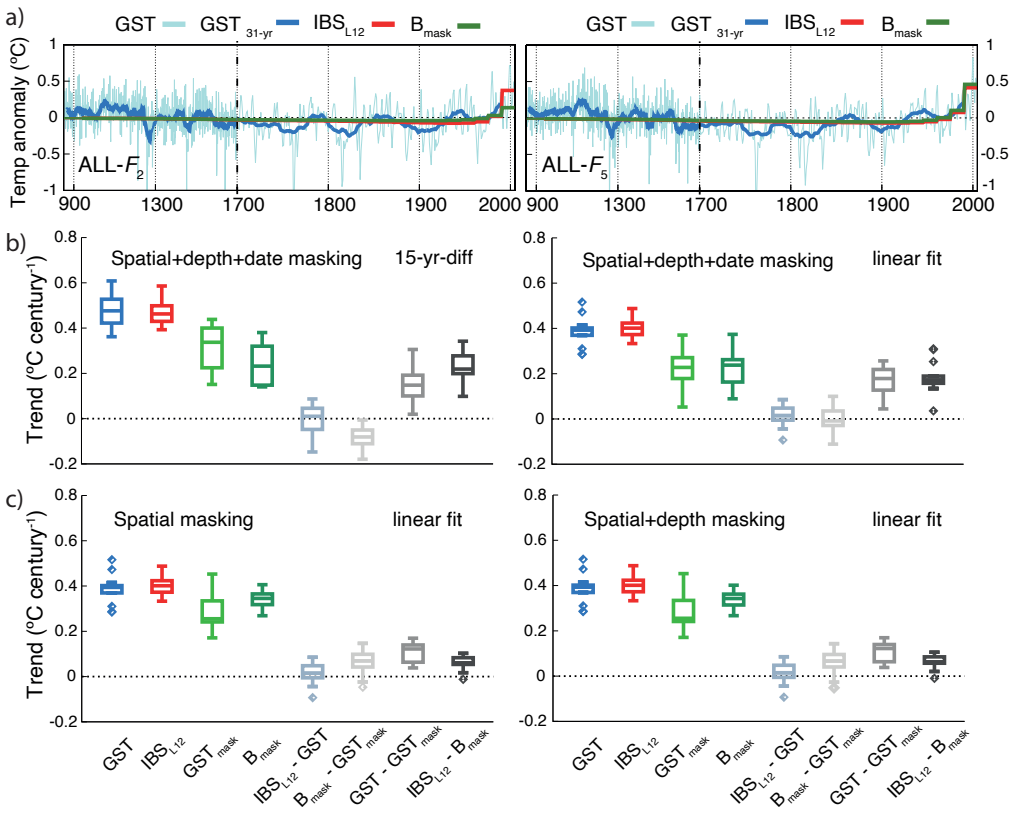

Figure 5.2: (a) LM global GST annual anomalies and the corresponding 31-yr filtered output, the global IBS_{L12} and the B_{mask} pseudo-reconstructions for the ALL-F_2 and ALL-F_5 members of the ALL-F ensemble. Note the different discretization in the x axis after 1700 CE. (b) Boxplots describing the centennial trends over the period 1900–2005 CE calculated for each of the 13 ensemble members within the ALL-F after applying space, depth and time masking in the model to mimic real BTP distribution. (c) as in (b) but applying spatial and depth masking (right) and spatial only masking (left). The GST, IBS_{L12}, GST_{mask}, B_{mask} and the differences between IBS_{L12}–GST, B_{mask}–GST_{mask}, GST–GST_{mask} and IBS_{L12}–B_{mask} cases are represented. Trends are presented as the 15-yr-diff (b left; see text) and the linear fit to the data calculated over an annual basis (b right). The 25th, 50th and 75th percentiles are indicated and the whiskers represent the lowest/highest value within 1.5 interquartile range (IQR) of the 25th/75th percentile. Outliers are indicated as diamonds in the same color of the box. They represent the values lower/higher than the lower/upper whisker. Note that colors in (a) correspond with those in (b) and (c)

is based on a t test and accounts for the temporal autocorrelation by using a lag-1 autoregressive statistical model, based on standard procedures for temperature time series (Santer et al., 2008; Hartmann et al., 2013). Likewise, autocorrelation is also accounted for in the estimation of the reduced degrees of freedom (von Storch and Zwiers, 1999). The trend values are somewhat smaller than those of the observational record, which range between 0.75 and 0.83 °C century $^{-1}$ over the 1901–2012 CE period for the global mean surface temperature (Hartmann et al., 2013). In both cases, the IBS_{L12} pseudo-reconstruction yields a reasonable estimation of the global GST increase during the 20th century. Note that the GST trends are slightly larger for the 15-year-difference case (15-year-diff) relative to the linear fit. While the linear trend indicates a median GST increase of 0.39, the 15-year-diff suggests a median of 0.47 °C century^{-1}. Nevertheless, the IBS_{L12}-GST differences are distributed around zero in either case (Figure 5.2b). The comparison of the two methods to estimate the 20th-century temperature increase shows that in either case, the pseudo-reconstructed GSTs robustly represent the targeted temperature signal. It is remarkable that when GSTs are masked, i.e., sampled, in space and time (GST_{mask}) following the real distribution, trends take a smaller range of values than those of GST. GST-GST_{mask} differences are above 0 and can reach 0.3 °C century^{-1}; thus they are significant and important in the context of the simulated warming. The significance test ($p < 0.05$) of trend differences is based on a "paired trends" test following Santer et al. (2000) and also accounts for temporal autocorrelation. It is remarkable how the level of impact may depend on the interplay of internal variability and BTP sampling. Additionally, the pseudo-reconstructions B_{mask}correspondingly deliver a distribution of trends that agrees in range with those of GST_{mask}. In the case of 15-year-diff, B_{mask} underestimates slightly, with B_{mask}–GST_{mask} distributed slightly below 0, whereas in the linear case B_{mask}–GST_{mask} differences are centered at zero.

Therefore, the borehole reconstructions are able to retrieve the masked or unmasked GST. However, sampling plays a role and can produce an underestimation of GSTs. This underestimation depends on the interplay between sampling and internal variability and occurs in some simulations in which for the selected points, their depths and logging dates are not optimal to represent global warming. Note that in Figure 5.2b B_{mask} actually includes the effects of masking, i.e., selecting, pseudo-BTPs co-located with the actual BTP network, trimmed to their actual depths in reality and generated using GST histories up to their logging dates. In turn, GST_{mask} includes the effects of masking, i.e., selecting in this case the grid-point time series co-located with actual BTP locations and trimmed to their logging dates; depth masking does not play a role in the case of GST series.

At this point, the question of what the relative role of each of the three masking effects is remains. Figure 5.2c addresses this by showing plots similar to the linear fit in Figure 5.2b but considering only spatial and spatial plus depth masking. Note that even if the masked version is still referenced only as $GST_{mask}(B_{mask})$ in the x axis, in Figure 5.2c left(right) the masking includes only spatial (spatial+depth) masking. The results for GST, IBS_{L12}, and IBS_{L12}-GST differences are shown in a way that is identical to Figure 5.2b (right) , for comparison. Results are virtually the same for both plots in Figure 5.2c. This implies that the effects of depth masking are negligible. Nonetheless, this may not be the case in real-world BTPs of different depths because the elimination of the background geothermal gradient to obtain the anomaly profiles is done by linear fitting at the bottom of the profile (Beltrami et al., 2011, 2015) thus introducing a source of uncertainty that has not yet been considered in PPE approaches. Spatial masking does have an impact, albeit smaller than if the effects of logging dates are considered, with GST–GST_{mask} differences above 0 and below 0.2 °C century^{-1}. IBS_{L12} and B_{mask} are effective in retaining their targets.

The selection of the number of singular values to be retained in SVD inversion may also exert some influence on the estimation of the GST recent trends. Therefore, the impact on reconstructed GST trends of including four and five PCs has been analyzed (not shown). On the one hand, the solution considering the four leading modes yields similar estimations as the solution based on the three leading modes. Nonetheless, the four-PC IBS_{L12} solution is slightly biased toward larger values. Such behavior is also present in the B_{mask} configuration relative to the GST_{mask}. This pattern is systematically observed in all members of the ALL-*F* ensemble. On the other hand, the results including the five leading modes yield a solution with some increase in the variance (Beltrami and Mareschal, 1995), noticeable within the 20th century. The simulated GST 20th-century trends are biased toward smaller values in comparison to those in Figure 5.2b; they are systematically underestimated by the five-PC IBS_{L12} solution. In real-world data, the level of noise in BTPs limits the number of retained principal components (Beltrami and Bourlon, 2004). We have used here a conservative low value that may be consistent with real-world applications including noise.

The hemispheric to global borehole reconstructions in the real-world conditions are far from the idealized scenario described by IBS_{L12} since they are the result of an aggregation of a limited number of BTPs that are sparsely distributed over the land surface. Further, there is a large variability in both depth and timing in the records since BTPs are obtained from different sources, and they are rarely drilled for the development of climate studies (Jaume-Santero et al., 2016).

The global GST increase during industrial times is, however, underestimated by the B_{mask} scenario. This feature is common to all ensemble members within the

ALL-F pool of simulations given the median IBS_{L12}–B_{mask} difference of 0.17 °C century^{-1} (Figure 5.2b), which accounts for about 43% of the simulated global GST 20th-century warming. Most of this underestimation is due to temporal masking. Almost half of the grid points containing BTPs in the B_{mask} case are dated prior to 1982 CE, and only ∼5% of them present logging dates after 1995 CE (Figure 5.1a). As a consequence, many of the synthetic BTPs do not include information from the last 2 decades of the simulated period, leading to a muted estimation of the global GST recent trend (Pollack and Smerdon, 2004). The variability in the depth of the borehole records, on the other hand, has no influence on the results of the B_{mask} configuration. This is because only BTPs deeper than 200 m have been retained. This is a depth that has been shown to be sufficient to capture the trend estimates from the LIA to present days and to even retain some features of the MCA to the LIA transition. Likewise, the spatial distribution of the borehole records has a smaller influence on the B_{mask} trends at hemispheric to global scales, as indicated in both real-world borehole reconstructions (Pollack and Smerdon, 2004; Beltrami and Bourlon, 2004) and PPEs (González-Rouco et al., 2006).

Despite the general pattern of discrepancies between B_{mask} and IBS_{L12} during the industrial period (e.g., ALL-F_2; Figure 5.2a), there are also cases in which both estimates show a good agreement (e.g., ALL-F_5). This range of variability in the representation of the 20th-century trends indicates that the internal climate variability within each realization exerts some influence on the simulated trends of the B_{mask} configuration. As the global average for the last 2 decades highly depends on the most recently logged BTPs, mostly concentrated in northern Canada, central Europe, Russia and Japan (Figure 5.1a), the internal climatic responses over these regions apparently play a significant role in the global 20th-century trend estimations. Nevertheless, it would be desirable to expand and update the current dataset of BTPs. In the meantime, it seems advisable to derive strategies that minimize the impact of aging in BTPs (Harris and Chapman, 2001), perhaps by blending BTP information and instrumental temperature observations from the last few decades or using other procedures that reduce underestimation of multi-decadal trends during this period.

The variability in the spatiotemporal distribution of the borehole network may impact the representation of the past GST evolution from BTPs at smaller spatial scales. In order to explore this, three subcontinental domains with different levels of BTP coverage have been selected. These domains include regions over North America, Europe and Africa. Their selection intends to be illustrative of sampling effects by broadly including existing BTPs in each domain without, however, necessarily representing optimized domain configurations in any sense.

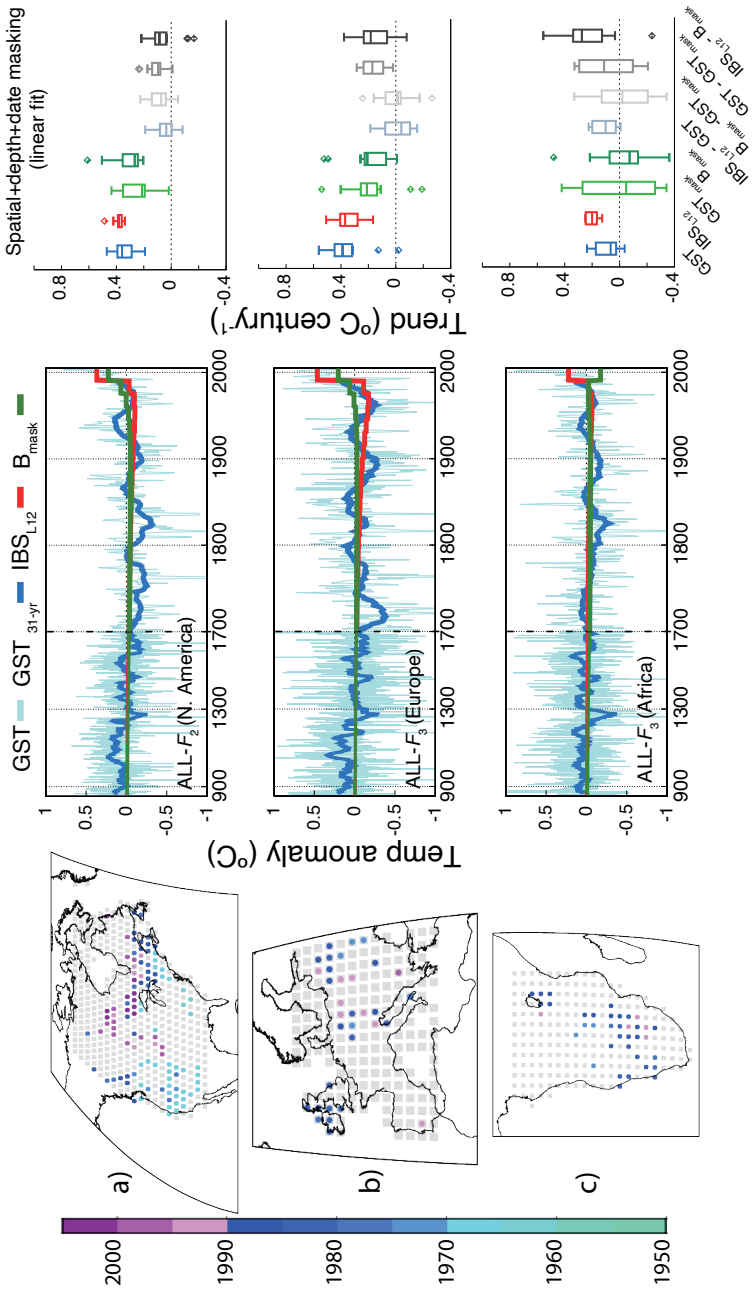

Figure 5.3: LM global GST annual anomalies and pseudo-reconstructions for three different sub-continental regions: (a) North America (31.26-69.16° N, 135-55° W), (b) Europe (36.59-59.68° N, 10° W - 30° E) and (c) Africa (35.05 ° S - 2.28° N, 10-40° E). The ALL-F_2 member is displayed for the North American continent, whereas the ALL-F_3 member of the ALL-Fensemble is used for the European and African regions. Maps in (a)–(c) show the spatial distribution of BTPs locations and dates of the actual borehole network for each of the regions. Grey dots show the model grid and the areas defining each of the regions. Right panels: as in Figure 5.2a and b but with the 1900–2005 CE trends presented only for the linear fit method.

The GST annual anomalies for these regions and their corresponding 31-year (running mean) low-pass filter outputs, IBS_{L12} and B_{mask}, estimates cases are shown for one member of the ALL-F ensemble as an example for each of the regions (Figure 5.3 center). ALL-F_2 is presented for North America (Figure 5.3a) while ALL-F_3 is shown for both Europe (Figure 5.3b) and Africa (Figure 5.3c) as these members are representative of the mean behavior of the B_{mask} configuration within the ALL-F ensemble. The box-and-whisker plots (Figure 5.3 right) show the 20th-century trends from the 13 members of the ensemble as in Figure 5.2b, using the linear regression.

Interestingly, the results of the IBS_{L12} and the B_{mask} configurations show better agreement over the better-sampled area of North America than over Europe and Africa. This is evident in the representation of the 20th-century trends, with a difference in their corresponding median value of around 0.08 °C century^{-1} in North America. Masking produces differences (GST–GST_{mask} and IBS_{L12}–B_{mask} in Figure 5.3) that range from 0 to 0.2 °C century^{-1}, thus suggesting underestimation biases. The bias is reduced when spatial-only masking is considered (Figure 5.4), although some outliers still produce GST–GST_{mask} differences above 0.1 °C century^{-1}. The size of those differences is not large, but it represents between 20% and 30% of the simulated trends.

The results for Europe and Africa are more variable. GST trends center around 0.4 °C century^{-1} in North America and Europe and below 0.1 °C century^{-1} in the African domain considered (Figure 5.3). However, trends at these spatial scales can be highly dependent on internal variability, and some simulations show no significant trends over the European and African domains. Additionally, poor spatial sampling enhances the influence of local behavior since only few grid points, often distributed within a relatively small area, determine the average of the whole region. Note that the representation of the 20th-century trends in both GST_{mask} and the B_{mask} spreads over a larger range in Europe and especially in the African region compared to North America. This happens as a response to a decline in the availability of recent BTPs for calculating the regional averages during the last few decades of the simulated period (see the maps in Figure 5.3). In Europe, the B_{mask} configuration systematically underestimates the IBS_{L12} results, but there are cases (not shown) in which the results of the B_{mask} case closely match the ideal scenario. Thus, these differences among estimates from the different ensemble members suggest an influence from the internal variability on the estimations of the recent temperature trends. Over the African continent, the estimates are more diverse. Therein, the difference between IBS_{L12} and B_{mask} depicts a larger variability, with a general poor representation of the 20th-century GST increase over this region.

When only spatial masking is considered, the spread of the results is reduced for all regions. For the European domain, most of the solutions of GST–GST_{mask} (IBS_{L12}–B_{mask}) get confined below +0.1 (-0.1) °C century^{-1} (Figure 5.4). For the African domain, the dispersion in GST–GST_{mask} and IBS_{L12}–B_{mask} also shrinks, with both cases indicating overestimation that can be larger than 0.3 °C century^{-1} when spatial masking only (Figure 5.4) is considered; i.e., the selected grid points indicate larger warming than the rest of the regions. The selected case examples in Figure 5.3 (center) illustrate how different the unmasked IBS_{L12} and the B_{mask} solutions can be, both in representing the warming of the last few decades of the 20th century and also the cooling during the 19th and early 20th century. The latter is most noticeable in the example shown for Europe in which clear differences develop during the 19th century between IBS_{L12} and B_{mask}. Here, spatial sampling misses the cooling shown by IBS_{L12}. The reasons for that are discussed in Section 5.2.2.

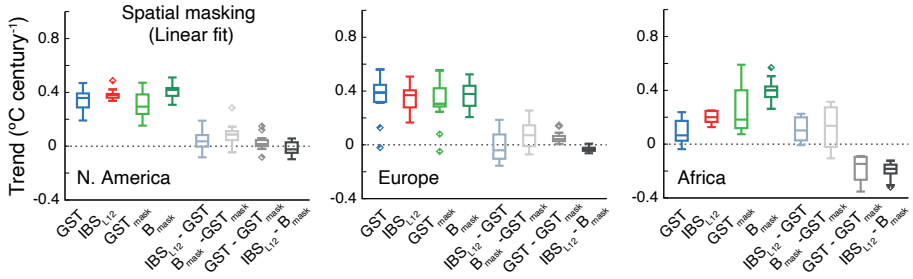

Figure 5.4: As in Figure 5.3 (right) but with the masked cases (i.e., GST_{mask} or B_{mask}) including only spatial masking. Linear fit estimated 1900–2005 CE trends in the ALL-*F* ensemble for North America, European and African regions shown in Figure 5.3. Abbrebviations for GST and PPE cases and their differences follow the same convention as in Figures 5.2 and 5.3.

Therefore, the 20th-century B_{mask} trends are sensitive to the spatiotemporal distribution of the borehole network, particularly at smaller scales. Whereas in the North American continent there is a better coverage of borehole logs, including most of those recently recorded, in Europe and especially in Africa, there are comparatively poorer representations and inhomogeneous spatial distributions of BTPs, particularly of the recent ones, that ultimately impact the resulting temperature trend estimates. This suggests that the interpretation of the trends from borehole reconstruction estimations at the regional scale should

be done with caution over areas with poor spatiotemporal coverage of BTPs (Huang et al., 2000).

5.2.2 Biases of past SAT borehole-based reconstructions due to physical processes

In addition to the limitations imposed by the spatiotemporal borehole distribution, the reconstructed SAT can also be biased by changes in the long-term SAT–GST relationship. Specifically, changes in anthropogenic forcing after 1850 CE can contribute to SAT–GST decoupling (Melo-Aguilar et al., 2018).

Figure 5.5a–c show the LM evolution of SAT and GST anomalies relative to the 850–2005 CE mean, as well as the difference between them, and the linear trend of SAT–GST for the period 1850–2005 CE at selected grid points, illustrating three possible behaviors of the long-term SAT–GST relationship in the CESM-LME. First, there is a case in which this relationship remains stable during the whole LM (Figure 5.5a). Note that there is no trend in the SAT-GST differences. This case represents the ideal strong SAT–GST coupling situation from which the GST would constitute a good proxy of the SAT. Second, two different cases in which the SAT–GST long-term relationship experiences significant variations are depicted. The direction and magnitude of the SAT minus GST trend are suggestive of the type of impact had on the SAT–GST coupling. In the first case (Figure 5.5b), the sharp decrease in GST around 1900 CE results in a positive trend in the SAT–GST differences of 1.04 °C century^{-1}, which is represented in red, indicating a warmer SAT relative to the GST. The second one (Figure 5.5c) depicts an example in which the SAT tends to be colder than GST during the industrial period, thus leading to a negative trend of about –0.4 °C century^{-1}. These two cases are suggestive of a physical interference affecting the temperature signal contained in the subsurface. Therefore, the inversion of BTPs under such characteristics would yield unreliable information of the past SAT variations. The ALL-F_2 simulation is used as an example, but similar results can be found in other ALL-F ensemble members.

To provide a spatial view of the borehole locations affected by changes in the long-term SAT–GST relationship in the CESM-LME world, we evaluate the SAT minus GST industrial (1850–2005 CE) trends for the ALL-F ensemble mean at every grid point with the presence of a borehole (Figure 5.5d). The ensemble mean is used in order to identify the forced response; however, the internal climate variability in individual ensemble members may result in a different representation of both the magnitude and sign of the trends (Melo-Aguilar et al.,

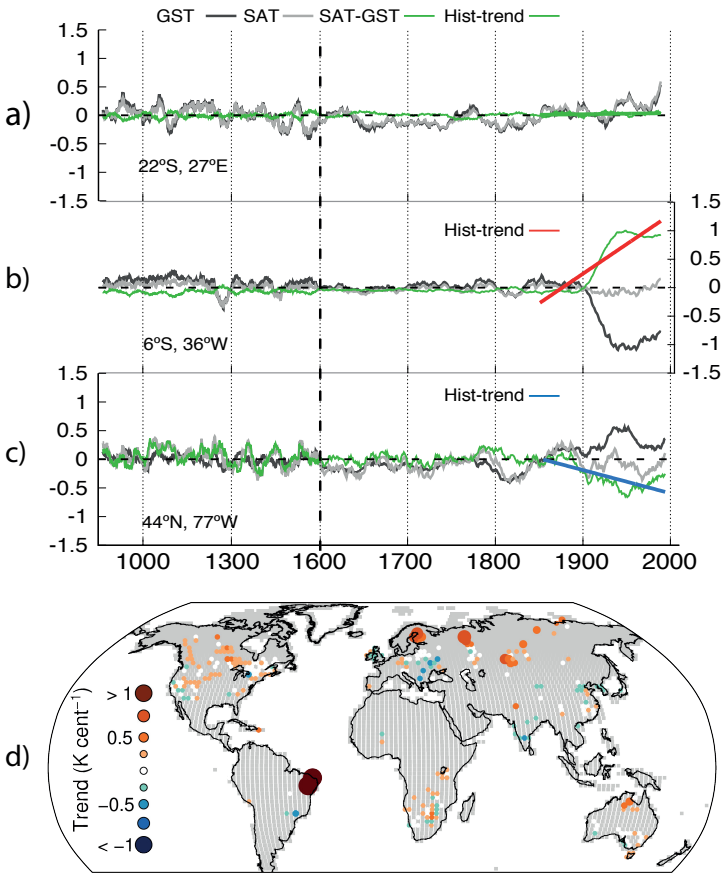

Figure 5.5: LM evolution of SAT and GST anomalies, SAT minus GST and the linear trends of SAT minus GST over the 1850–2005 CE period for three different grid points with borehole records in the ALL-F_2 as an example. Each grid point illustrates a possible case of the long-term SAT–GST relationship: (a) strong coupling during the LM, (b) decreasing of GST relative to SAT and (c) increasing of GST relative to SAT during industrial times. d) Spatial distribution of linear trends in SAT minus GST anomalies over the 1850–2005 CE period evaluated at every grid point with the presence of a borehole site using the ALL-F ensemble mean. Grid points showing statistically significant trends ($p < 0.05$) are colored in red/blue depending on whether they are increasing/decreasing SAT–GST trends.

2018). Since the industrial period is affected by pronounced temperature trends due to the anthropogenic emissions of GHGs, the 1850–2005 CE interval is the most adequate one to use for evaluating the particularities of the SAT–GST linear trends. Additionally, the analysis in this part is intended to determine the reliability of the borehole technique for retrieving the SAT increase over the industrial period. Figure 5.5d illustrates that at a large number of grid points containing BTPs, the SAT minus GST linear trends are statistically significant ($p < 0.05$), identifying therefore some level of SAT–GST decoupling during the industrial period at such locations. These grid points are represented in red (blue) if the trend of the temperature differences is positive (negative). Additionally, the sizes of the circles are related to the magnitude of the trend. Therefore, both color and size indicate the direction and magnitude of the SAT–GST decoupling. Note that positive trends are dominant, with some larger values in northeastern Brazil, Fennoscandia, central Eurasia and northern Siberia. Negative values are also evident, mostly distributed around central and eastern Europe, with some more isolated cases distributed around the globe.

In order to assess the influence of the detected long-term SAT–GST decoupling on the representation of SAT from the borehole-based reconstructions at the global scale, results of the SAT and GST trends and their IBS pseudo-reconstructions, as well as the B_{mask} case including the effects of sampling, are shown in Figure 5.6, together with the ALL-F_2 and ALL-F_5 ensemble members as an example. In both of them, the simulated SAT low-frequency variations over the LM are broadly reproduced by IBS_{SAT}. Note that the simulated SAT 20th-century trend is accurately captured by the IBS_{SAT} since the differences between ensemble member trends are distributed around zero (Figure 5.6b). Therefore, IBS_{SAT} can be considered a reasonable representation of the simulated SAT, specifically in the estimation of the 20th-century warming. Thus, the comparison between IBS_{SAT} and IBS_{L12} will be informative about any deviation between the pseudo-reconstructed GST and the simulated SAT.

Indeed, SAT–GST trend differences are distributed over a median value of 0.1 °C century^{-1}, and consistently, the 20th-century IBS_{SAT}–IBS_{L12} trend differences yield positive values, with a median of about 0.11 °C century^{-1} and a high level of agreement, i.e., small dispersion, among the members of the ALL-F ensemble (Figure 5.6b). This indicates that even under an ideal scenario, the pseudo-reconstructed GST does not fully capture the 20th-century SAT warming, missing on average ∼20% of the simulated SAT increase in the CESM-LME as a response to the long-term SAT–GST decoupling. Furthermore, the effect of the physical processes is superimposed on the limitations due to the methodological aspects, leading to larger differences with respect to the more realistic pseudo-reconstructed GST (B_{mask}). In fact, the IBS_{SAT}–B_{mask} 20th-century trend dif-

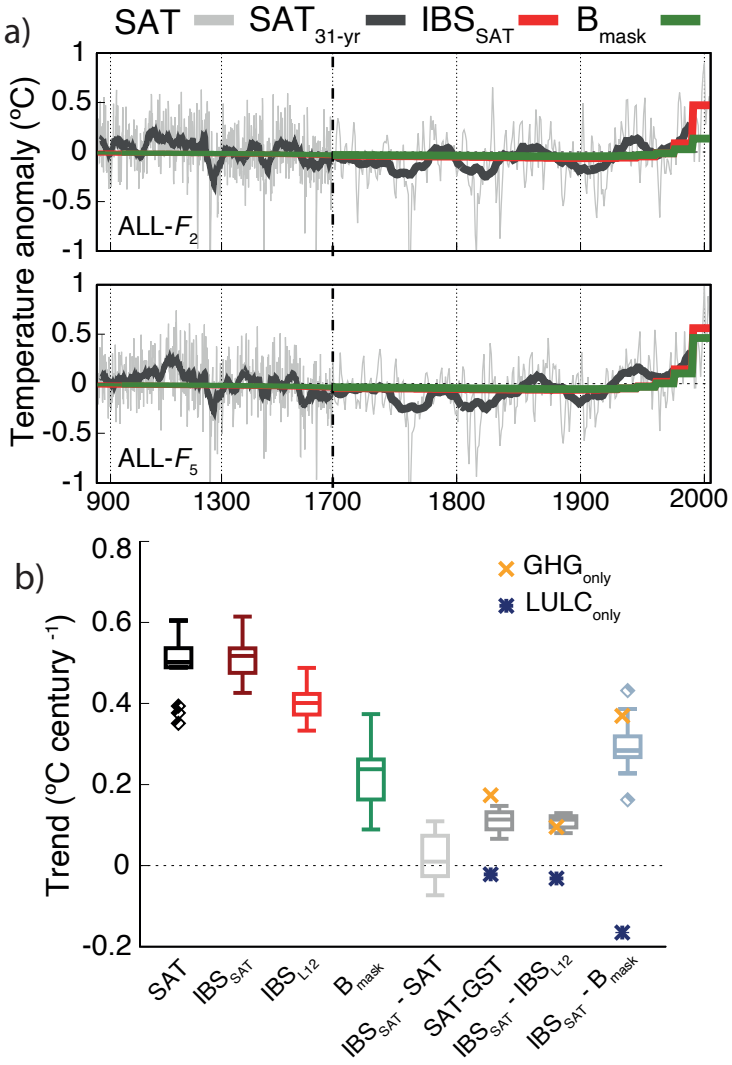

Figure 5.6: (a) LM global SAT annual anomalies and the corresponding 31-yr filtered outputs and the global IBS_{SAT} and B_{mask} pseudo-reconstructions for the ALL-F_2 and ALL-F_5 members of the ALL-F ensemble. Note the different discretization in the x axis after 1700 CE. (b) Estimated linear fit as Figigure 5.2b, but the SAT, IBS_{SAT}, IBS_{L12}, B_{mask} and the differences between IBS_{SAT} - SAT, SAT–GST, IBS_{SAT}–IBS_{L12} and IBS_{SAT}–B_{mask} cases are represented. Additionally, the ensemble median of the GHG- and LULC-only ensembles is indicated by the crosses and asterisks, respectively, in the IBS_{SAT} - IBS_{L12} and the IBS_{SAT}–B_{mask} columns.

ferences have a median of 0.28 °C century^{-1} (Figure 5.6b), which represents more than 50% of the simulated SAT trend. There is, however, a larger variability in trend differences between these two scenarios, which ranges between 0.16 and 0.43 °C century^{-1}. This spread results from the internal climate variability in each realization of the ensemble. Such a range of variability is noticeable in Figure 5.6a if IBS_{SAT} and B_{mask} trends for the two ensemble members are compared to each other.

Melo-Aguilar et al. (2018) showed that the long-term SAT–GST decoupling in the CESM-LME is mainly driven by LULC and GHG changes. LULC changes modify the radiative fluxes at the surface, leading to a different response of the SAT and GST. Likewise, the SAT increase during industrial times as a response to the increase in GHGs may not be fully transferred to the soil due to the insulating effect of snow cover feedbacks. To explore their effects on borehole reconstructions, the analysis described in Figure 5.5d is extended to the GHG- and LULC-only ensembles (Figure 5.7a and b, respectively). In the GHG-only ensemble the dominant effect is represented by positive SAT minus GST trends (Figure 5.7a), with a similar pattern to the ALL-*F* ensemble over North America, Fennoscandia, northern Russia and Siberia in Figure 5.5d, although the magnitude of the trends is significantly larger. On the contrary, in the LULC-only ensemble negative SAT minus GST estimates are evident (Figure 5.7b), indicating a similar response as in the ALL-F ensemble over central and eastern Europe and the Indian subcontinent (Figure 5.5). Additionally, the strong positive trends over northeastern Brazil resemble those found in the ALL-*F* ensemble.

The IBS_{SAT}, the IBS_{L12} and the B_{mask}configurations are also implemented in the GHG- and LULC-only ensembles in order to evaluate their contributions to the physical SAT–GST decoupling at the global scale. The bottom panels in Figure 5.7a,b compare the simulated global SAT anomalies with the IBS_{SAT} and the B_{mask} pseudo-reconstructions in the GHG_1 and $LULC_1$members of the GHG- and LULC-only ensembles, respectively, as an example. From a simple visual inspection it may be noticed that the GHG_1 simulation portrays a response similar to the ALL-*F* estimates (Figure 5.6a) since the IBS_{SAT} and the B_{mask} cases diverge during the last few decades of the simulated period. On the other hand, in the $LULC_1$ case, there is no such similarity to the ALL-*F* cases. All series and pseudo-reconstructions suggest a consistent cooling in response to LULC changes throughout the last few centuries of the millennium. This analysis suggests a larger contribution of the GHG forcing to the SAT–GST decoupling relative to the LULC forcing at the global scale.

To provide a quantitative support for this statement, the median of the differences between the IBS_{SAT}–IBS_{L12} and IBS_{SAT}–B_{mask} 20th-century trends in the GHG- and LULC-only ensemble are included in Figure 5.6b (crosses and

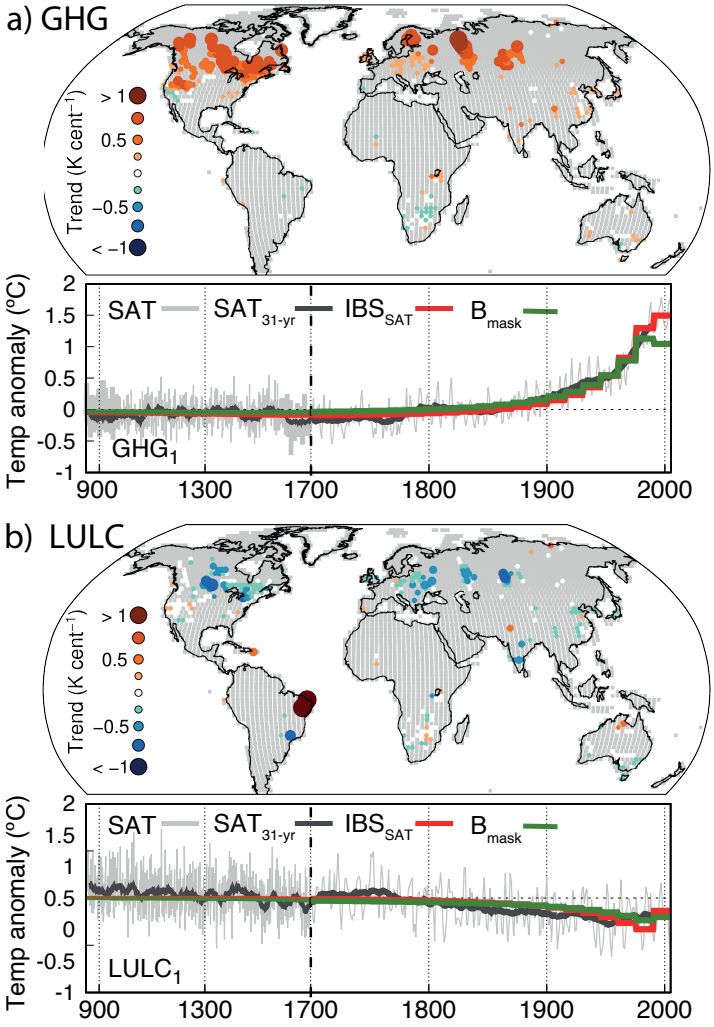

Figure 5.7: (a) Spatial distribution of linear trends in SAT–GST anomalies over the 1850–2005 CE period evaluated at every grid point co-located to a borehole site in the GHG-only ensemble mean. Only grid points delivering statistically significant trends ($p < 0.05$) are colored. The bottom panel shows the LM global SAT annual anomalies and the corresponding 31-yr filtered outputs, the global IBS_{SAT} and the B_{mask} pseudo-reconstructions for the GHG_1 and $LULC_1$ ensemble members. Note the different discretization in the x axis after 1700 CE. (b) as in (a) but for the LULC-only ensemble.

asterisks, respectively). Results are almost identical if the mean instead of the median of the single-F is included. Note there that in the IBS_{SAT}–IBS_{L12} case the GHG-only ensemble shows positive and very similar estimates to the median of the ALL-F ensemble (0.09 and 0.11 °C century^{-1}, respectively), suggesting a large contribution of the GHG forcing to the physical bias in the representation of SAT by the borehole pseudo-reconstructions in the CESM-LME at the global scale. On the contrary, in the LULC-only ensemble the IBS_{SAT}–IBS_{L12} difference is negative and very small (–0.02 °C century^{-1}) indicating a negligible contribution at this spatial scale. These results are in agreement, as expected, with pure SAT–GST differences, which also receive a larger contribution from GHG at these spatial scales. The IBS_{SAT}–B_{mask} differences further highlight the larger influence of the GHG relative to the LULC forcing. Whereas in the GHG-only ensemble the median IBS_{SAT}–B_{mask} difference shows values trending in the same direction as the ALL-F ensemble, in the LULC-only ensemble the difference goes in the opposite direction (negative IBS_{SAT}–B_{mask} trend differences). Melo-Aguilar et al. (2018) reported that the contribution from the GHG forcing to the SAT–GST decoupling is controlled by the reduction in the NH winter snow cover as a response to higher temperatures during industrial times. This situation increases the exposure of the soil surface, previously insulated by snow cover, to the cold winter air, leading to an overall effect of warmer SAT relative to GST at a global scale.

Regionally, the influence of the SAT–GST long-term decoupling on the representation of simulated SAT from the pseudo-reconstructed GST also deserves to be considered since there are geographical variations in this effect (Figure 5.5d). Figure 5.8 illustrates SAT annual anomalies with respect to the 850–2005 CE mean and the corresponding 31-year low-pass filter outputs as well as the IBS_{SAT} and the B_{mask} cases for the ALL-F_2, GHG_1 and $LULC_1$ simulations over North America (a), Europe (b) and Africa (c) including the same spatial domains depicted in Figure 5.3 for these three regions. The spreads provided by considering all members of each ensemble are depicted in the boxplots at the bottom of Figure 5.8.

Interestingly, the decoupling effect appears to be larger over North America and Africa than over Europe. As in the case of the global average, the effect over North America and Africa is shown by the underestimation of the SAT warming during industrial times, indicated by the positive SAT–GST median differences of ∼0.2 °C century^{-1}, consistent with comparable ∼0.2 °C century^{-1} median change in IBS_{SAT} - IBS_{L12} that endorse the performance of the borehole method. On the contrary, in Europe, the SAT–GST decoupling leads to small differences, a negligible under- or overestimation of the SAT increase as shown by SAT–GST (IBS_{SAT}–IBS_{L12}).

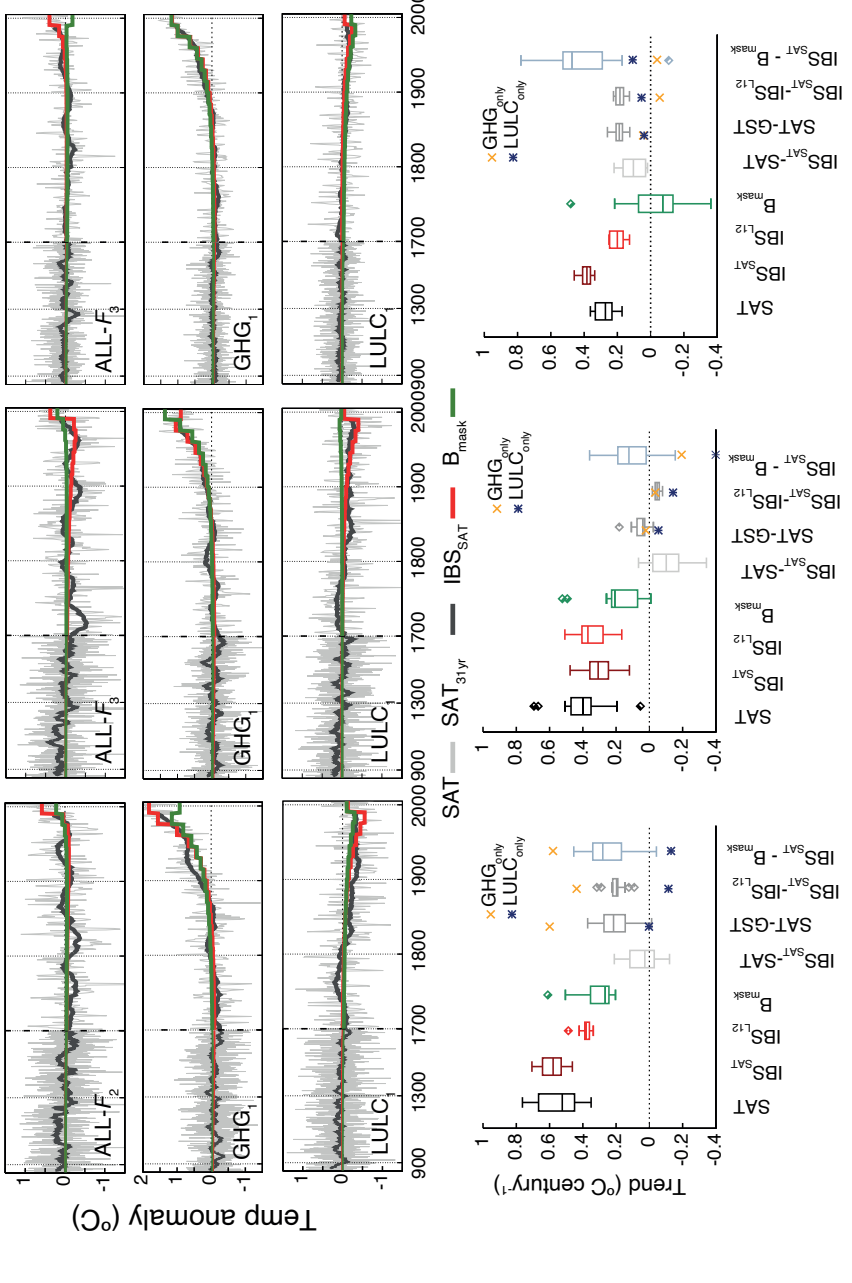

Figure 5.8: LM regional SAT annual anomalies and the corresponding 31-yr filtered outputs, the global IBS_{SAT} and the B_{mask} pseudo-reconstructions for the ALL-F_2, GHG_1 and $LULC_1$ (from top to bottom) members of the ALL-F, GHG and LULC-only ensembles, respectively: North America (a), Europe (b) and Africa (c). Note the different discretization in the x axis after 1700 CE. Bottom panels: as Figure 5.6b but for each of the regions presented.

The single-F experiments indicate variability in the influence of both the GHG and LULC forcings over the different areas considered herein. In North America there is a strong contribution of the GHG forcing, which is somewhat counteracted by the LULC influence. This is noticeable in not only the time series in Figure 5.8a but also the sign and magnitude of the SAT–GST and the IBS_{SAT}–IBS_{L12} median difference in both the GHG- and LULC-only ensembles (Figure 5.8a bottom). Conversely, over the European region, the largest contribution comes from the LULC forcing (see the larger differences in SAT – GST and IBS_{SAT}–IBS_{L12} for LULC-only in Figure 5.8b bottom). This is shown more clearly in the comparison of the selected examples in Figure 5.8b for ALL-F_2, GHG_1 and $LULC_1$ that show that the LULC cooling dominates the long-term trends over GHG warming in the ALL-F experiment. For Africa, the ALL-F experiment also evidences the damping of the GHG warming produced by negative LULC trends. However, other external forcings may also contribute to the response of the SAT–GST relationship at the continental scale in this region (Melo-Aguilar et al., 2018). It is also remarkable that the masked reconstruction (B_{mask}) in the ALL-F and LULC cases in Figure 5.8b does not capture the long-term cooling during the 19th and 20th centuries that is shown by the unmasked SAT averages (Figure 5.8b) and GST averages (Figure 5.3b center). Thus, the influence of LULC forcing also explains the different trends between IBS_{L12} and B_{mask} in Figure 5.3b (center) during the 19th and early 20th centuries.

The superposition of the physical biases and the methodological aspects at continental scales suggest further comments are needed. For instance, for the North American region, while the methodological constraints have a relatively reduced impact on retrieving the simulated GST 20th-century increase in the B_{mask} configuration (as discussed in Section 5.2.1), the effect of the physical processes results in a larger underestimation of the 20th-century SAT evolution. This is evident in the IBS_{SAT}–B_{mask} difference, which has a median of 0.28 °C $century^{-1}$ (Figure 5.8), representing ∼50% of the simulated SAT 20th-century warming. The increment can be compared with IBS_{L12}–B_{mask} in Figure 5.3 (right). The largest contribution comes from the GHG-only forcing, partially reduced by the LULC-only effect, as shown by the single-forcing crosses in Figure 5.8. Similarly, in the African region, the effects of the physical processes contribute to enhancement of the underestimation of the SAT signal by B_{mask}. IBS_{SAT}–B_{mask} differences expand their spread in Figure 5.3, now reaching values of 0.8 °C $century^{-1}$. On the contrary, for the European region the SAT–GST decoupling slightly counteracts the effects of the methodological aspects, mostly via negative LULC biases. However, in both of these regions, the largest contribution to a biased estimation of the 20th-century SAT trends comes from the methodological aspects (see

SAT–GST and IBS_{L12}–B_{mask} differences in Figure 5.3) that outweigh the bias from the physical processes.

5.3 Conclusions

Borehole-based reconstructions can be affected by a set of methodological constraints and some physical mechanisms that corrupt the SAT–GST coupling assumption. While the methodological issues may distort the recovery of past GSTs from BTPs, the physical processes may play a role in recovering past SAT changes from the reconstructed GST. This chapter has addressed the borehole temperature reconstruction technique's performance in a pseudo proxy framework that allows for addressing both methodological and physical issues.

Pseudo-proxy approaches to evaluate the borehole technique with a focus on the methodological performance have been implemented in simulations of the LM at global (González-Rouco et al., 2006; García-García et al., 2016) and regional scales (González-Rouco et al., 2009), either using the output of a single climate model (González-Rouco et al., 2006, 2009) or an ensemble of PMIP3/CMIP5 LM experiments (García-García et al., 2016). In the present chapter, these analyses have been extended by implementing a PPE strategy within the ensemble of simulations of the LM produced with the CESM-LME. Thus, an update of previous works is presented by introducing a more realistic PPE setup to address methodological and physical issues at global and regional scales. Furthermore, the use of an ensemble of simulations with the same model and different forcing boundary conditions has allowed for considering the influence of internal variability and external forcings on applying the borehole method.

Similar to previous studies (González-Rouco et al., 2006, 2009; García-García et al., 2016), the methodological implementation used herein adopts a standard SVD inversion approach. Likewise, the PPE has been developed in a so-called idealized scenario in which BTPs are assumed to exist at every land model grid point, and produced with a forward model using the complete simulated GSTs, 850-2005 CE. This is equivalent to considering that information from BTPs is available everywhere in land and with logging dates updated to present times. Also, more realistic scenarios considering distributions of BTPs that mimic their actual spatial, depth and logging date distributions have been constructed. The idealized scenario is used as a benchmark from which the performance of the realistic case is evaluated.

Regarding physical influences on the SAT–GST relationship, this chapter has been built on the results of Chapter 4 (Melo-Aguilar et al., 2018) that demon-

strate with this ensemble of CESM experiments that external forcings and related feedbacks (e.g., snow cover) can have an impact on SAT–GST coupling, with implications for borehole reconstructions, particularly at regional scales. The work developed herein implements a PPE setup and analysis on the same model ensemble. This allows a consideration of the SAT–GST changes in simulations, including a full configuration of natural and anthropogenic forcings and some single-forcing simulations that can help interpret the results. Differences between SAT and GST trends and between borehole reconstructions using SAT and GST generated BTPs allow an understanding of the limits of the long–term coupling hypothesis.

The CESM-LME has been reported to underestimate by 20% the 20th-century trends (Otto-Bliesner et al., 2016). In the ensemble of simulations including all natural and anthropogenic forcings (ALL-F ensemble), linear fit GST (SAT) trends vary among ensemble members ranging between 0.28 and 0.51 (0.35 and 0.60) °C century^{-1}. This inter-simulation variability is an expected result of changing the initial conditions to generate the ensemble and reflects internal variability. The estimation of 20th-century trends has been employed as a metric of comparison of simulated and pseudo-reconstructed evolution by considering two different approaches: (1) differences between the final and initial 15-year periods and (2) linear trends; both of them are consistent in delivering robust results.

In the so-called idealized IBS PPE setup, a similar range of pseudo-reconstructed GST trends as the simulations is observed, thus supporting the overall performance of the SVD technique. This method can retrieve the long-term trends through the LM and the warming during the industrial period. Thus, the SVD approach itself renders reliable results in terms of retrieving the boundary GST signal. Some sensitivity to an increasing number of SVD modes is evident, though. A number of modes consistent with previous modeling and experimental studies were selected here. Results are robust to small changes in this configuration.

A common result that affects all the subsequent tests in this assessment is that when methodological and physical constraints are imposed on the borehole reconstruction method, results depend on initial conditions and therefore on internal variability. This is a new although also arguably expected result, as imposing a specific spatial and temporal sampling setup at global and regional scales may produce different effects on 20th-century trends depending on the particular trajectory of internal variability and how effective the distribution of BTPs is in grasping the global and regional warming signals embedded within the range of internal variability. The results indicate that sampling can introduce detectable biases in borehole reconstructions both at global and regional scales. In the specific setup included herein, considering a realistic distribution of depths does not

produce any detectable impact. This may be, in part, due to the little signal contained in the synthetic BTPs below 200–300 m depth because the CESM-LME simulations depict relatively low multicentennial surface temperature variability. The distribution of logging dates and BTP locations, on the other hand, does have some impact. At global scales, spatial and temporal masking introduces biases that can range between >0 and 0.3 °C century^{-1} (GST–GST_{mask} and IBS_{L12}–B_{mask} differences); spatial-only masking reduces maximum differences to 0.1 °C century^{-1}. This means that some simulations experience no significant change, and others, masked PPE can experience underestimations of several tenths of a degree depending on the realization of internal variability. These are indeed small numbers but bear in mind that even a 0.1 °C century^{-1} change represents about 20% of the total warming in these simulations.

Regionally, temporal and spatial sampling impacts vary across regions, with differences (e.g., IBS_{L12}–B_{mask}) that can range from 0 to 0.2 °C century^{-1} in a relatively well-sampled region such as North America to 0.4 °C century^{-1} in Europe, or 0.6 in the African domain. Most of it is due to the temporal sampling effect, particularly in the American and European regions where it gets reduced to values below 0.2 and 0.1 °C century^{-1}, respectively, when considering spatial-only masking; in the African domain, spatial biases are larger and range between 0.1 and 0.3 °C century^{-1}. At regional scales, some of these biases are larger than their target temperature trend values. Thus, at these scales, spatial and temporal sampling is an issue. The effects are smaller over North America and larger over the other regions tested.

Temporal logging of the borehole records stands as the main sampling aspect contributing to the reduced skill in capturing the 20th-century warming. This is because many BTPs do not contain information about the warming during the last few decades of the 20th century due to their relatively old logging dates. Such a result suggests that recent measurements' availability highly influences the results of this type of temperature reconstruction. The continental-scale analysis has provided further insight into this issue since the pseudo-reconstructed GST yields a generally good estimation of the simulated GST evolution during industrial times for areas with a relatively good distribution of recent BTPs measurements, such as the North American continent. On the contrary, this accuracy is lost as the availability of recent BTPs is reduced, for instance, over the African region. The temporal effects should be relatively small if the reconstructions are considered only up to the dates of the oldest boreholes, at the expense of missing much of the warming developed during the last decades. Alternatively, strategies may be considered to blend information from early borehole profiles with local instrumental data to mitigate the missing trend effect (Harris and Chapman, 2001). In addition, this type of analysis would benefit from re-logging of boreholes

whenever possible and logging of additional BTPs in the future, thus updating the network. Regarding the spatial sampling effects, the definition of the domains considered herein has been somewhat subjective. Perhaps more ad hoc domain setups that reduce spatial sampling effects like in Africa can be specifically considered. Otherwise, cases like the one selected in the African domain, including very sparse sampling, should be avoided.

In the surrogate reality of the CESM-LME, the interpretation of the simulated SAT, derived from the reconstructed GST, is additionally impacted by the physical processes that interrupt the long-term SAT–GST coupling. In the idealized scenario, the GST pseudo-reconstruction does not fully capture the global SAT increase during industrial times, missing about 20% of the simulated warming on average in the ALL-F ensemble. Globally, the larger increase of SAT relative to GST during industrial times arises from the GHG external forcing's dominant influence. Nevertheless, the contribution of individual forcings varies geographically, as indicated by the regional analysis. While over the North American continent, the overall response is similar to that at the global scale, with a large contribution of the GHG forcing, over Europe and Africa, the SAT–GST decoupling is dominated by the LULC forcing.

The impact of the long-term SAT–GST decoupling is superimposed on the limitations due to the methodological aspects. On a global scale, the combined effect results in a biased representation of the simulated SAT 20th-century trend, with the realistic configuration, missing more than 50% of the simulated SAT (average values for the ensemble). Indeed, at this spatial scale, the largest contribution comes from the methodological aspects. Nonetheless, this may be different at smaller spatial scales. For instance, in North America, where the impact of the methodological aspects is relatively small, the SAT–GST decoupling deteriorates the representation of the simulated SAT from the pseudo-reconstructed GST.

Overall, the results indicate that the combined effect of the methodological constraints and the physical SAT–GST decoupling leads to a systematic underestimation of the 20th-century SAT trends globally; at regional scale, overestimations can occur as in the African domain. Nonetheless, it is worth noting that despite this general pattern among the members of the ALL-F ensemble, there may be cases in which these impacts are relatively small or even disappear, consistent with previous studies (González-Rouco et al., 2006; García-García et al., 2016). This influence of internal variability and comparison with the actual case in the real world can be better evaluated over reanalysis simulations of the 20th century (Hartmann et al., 2013), currently underway.

Even if the analysis included herein represents the most realistic implementation to date of the borehole method in PPEs, results are not meant to be directly translated to real-world cases. Nevertheless, PPEs provide valuable in-

formation about the uncertainties in paleo-reconstructions in a controlled experimental framework (Smerdon, 2012). Although the results in this chapter indicate a clear pattern in the potential sources of bias of the borehole-based reconstruction, they should be interpreted with caution for real-world applications since several aspects may influence the level of impact. For instance, the use of other ESMs with different climate sensitivities, i.e., larger 20th-century warming, and a different representation of the influence of changes in land surface physics could lead to a different representation of changes in the SAT–GST relationship. Additionally, the limitations in the local representation of sampling due to model resolution and more technical issues like the existence of local noise in BTPs have not been considered here.

Considering the results of this chapter, both the methodological issues and the borehole-based reconstructions physical biases would lead to underestimating the temperature increase from the LIA to the present day, especially over the industrial period. These findings cannot explain the larger temperature increase suggested by the actual borehole estimations during the last few centuries of the LM relative to those of some other proxy-based reconstructions.

6

Conclusions and outlook

This Thesis evaluates the coupling between the air and the ground surface temperature and the propagation of the surface temperature changes into the subsurface at different spatial and temporal scales. As these two aspects are central assumptions for the development of paleoclimatic analyses using present-day BTPs, this Thesis also explores the potential sources of bias that can impact the temperature reconstructions based on BTPs. Each chapter of the text contains a specific section devoted to summarizing the main findings and conclusions therein. However, a discussion of the general conclusions of this Thesis is provided in the following section, thus framing a broader perspective of the work.

6.1 Main conclusions

A soil temperature observational database for Spain, SoTOS, has been developed, consisting of 39 stations spanning the 1985–2018 period. From an initial number of more than 300 time series, the final database consists of 38 sites over peninsular Spain and one site in the Balearic Islands at daily time resolution. The dataset includes soil temperature observations at 0.05, 0.1, 0,2, 0.5 and 1 m depth, and 2m air temperature with a maximum time span of 33 years (1985–2018) where the last 10 years show a rather temporally stable coexistence of around 20 stations. Additionally, 20 sites are still active in 2018. SoTOS provides a range of opportunities for assessments on the soil thermal state for several scientific applications, including drought risk management, land-atmosphere feedback on climate, and

model evaluation analyses among others.

The conductive heat transport explains, to a large extent, the behavior of the shallow subsurface thermal regime in SoTOS. The soil layer temperatures included in SoTOS effectively illustrate the conductive-dominated heat transport within the subsurface of the propagating surface temperature signal. They support the linear dependency of either of the amplitude attenuation and the phase shift of the temperature at different soil layers with the depth. Thus, allowing to estimate the mean thermal diffusivity of the upper soil meter for 39 locations over Spain as a measure of the conductive heat transport within the subsurface.

There exist seasonal processes influencing the SAT–GST coupling as indicated by instrumental records in SoTOS, ERA5-L reanalysis and CESM-LME simulations. It is broadly recognized in the literature that the coupling between SAT and GST can be disturbed by different processes acting at seasonal timescales such as, for instance, the insulating effect of snow cover that prevents the transfer of energy between the ground surface and the atmosphere.Using SoTOS, this Thesis has shown that other seasonal processes may lead to different SAT and GST responses. This is the case of the radiative heating of the ground surface in the summer months that produces higher GST relative to the overlying air over several inland locations. The high incoming solar radiation may explain the nature of such a response in addition to the low soil moisture content in the summer months over continental Spain. Such a behavior is also represented in the ERA5-L reanalysis. The use of the CESM-LME simulation reveals further mechanisms that can lead to different responses in SAT and GST on seasonal timescales. For instance, enhancing latent heat fluxes in the summer season over some regions around the globe results in a larger SAT warming relative to GST due to the evaporative cooling of the ground surface. Similarly, the role of snow cover as the most important agent in modulating the connection between SAT and GST is supported by this set of simulations and by ERA5-L.

The SAT–GST coupling assumption can also be impacted on longer temporal scales. In addition to the different responses of SAT and GST at seasonal scales, processes acting at longer timescales can also impact the relationship between SAT and GST. The use of instrumental records has provided insights into some of the mechanisms that produce variations in the SAT–GST relationship at interannual to decadal timescales. In this case study, a larger temperature increase of the ground surface compared to the SAT is represented at some locations in Spain over the last few decades. Decreasing precipitation and surface humidity rates presumably drive such a response during this period. The latter leads to

enhanced radiative heating of the ground surface as less water is available to evaporate, producing a warmer GST relative to SAT. The CESM-LME simulations have broadened our understanding of the SAT and GST relationship since they provided a wider spatial view and longer time frame on this issue. On a global scale, the CESM-LME simulations have shown that the SAT–GST coupling is strong above multi-decadal timescales since the GST evolution closely tracks SAT throughout the LM. This supports the strong SAT–GST coupling assumption at these spatiotemporal scales. Nevertheless, the CESM-LME also reflects that the SAT–GST relationship has not remained constant along the whole LM since a relatively small offset between the simulated SAT and GST during industrial times is also evident. This effect is negligible globally as it only represents ca. 0.05 °C difference during the 20th century. However, at regional to local spatial scales, there are cases in which the long-term SAT–GST relationship may be disrupted significantly due to modifications in the surface energy balance, mostly over the industrial period.

The influence of external forcings represents the main driver for long-term decoupling between SAT and GST with potential implications for interpreting SAT from borehole GST reconstructions, particularly at regional scales. The Single-*F* simulations within the CESM-LME have shown that the influence of some of the LM external forcings may lead to long-term modifications of the energy fluxes at the surface either from increased evapotranspiration, reduced energy dissipation via sensible heat and others that are the main physical mechanisms controlling the SAT–GST decoupling at timescales relevant for the borehole theory. Such variations in the surface energy balance are mostly controlled by the response to LULC changes triggered by agriculture expansion, mainly during the industrial period. For instance, deforestation at mid and high latitudes leads to long-term changes in the SAT–GST relationship due to albedo, surface roughness, and hydrology variations. Over tropical regions, deforestation limits the energy dissipation from the ground surface via latent or sensible heat fluxes resulting in larger GST warming relative to SAT. Similarly, the interactions of various mechanisms and feedbacks from different external forcings can also influence the long-term SAT–GST relationship at different spatial scales. Specifically, increasing GHGs concentration during industrial times leads to a reduction of the snow cover counterbalanced by the opposite effect of both LULC and OZ/AER forcings. Accordingly, the snow cover remains relatively stable over some regions during the simulated industrial period in the presence of a warmer climate, leading to a difference in the SAT and GST long-term evolution during the industrial period over the NH high-latitudes since the snow cover insulates the soil from a warmer SAT. In areas where the SAT–GST relationship is impacted, the past

GST evolution retrieved from BTPs would no yield a reliable representation of the past SAT variations.

Methodological constraints can impact achieving a robust estimation of the GST 20th-century trends from Borehole-based reconstructions. The development of PPEs using the CESM-LME simulations has provided insights into the effects of methodological issues, such as the availability of BTPs around the globe, the dates when they were logged and the depth they reach. The pseudo-proxy strategy implemented in this Thesis shows that in the absence of these methodological issues, i.e., the so-called idealized IBS PPE setup, is indeed able to produce pseudo-reconstructed GST evolution over the LM in a similar range as the simulations. While this supports the overall performance of the borehole method, as the methodological aspects are gradually imposed, there is an impact on the recovered 20th-century GST trend. Particularly, the spatial distribution and actual logging dates of BTPs influence the pseudo-reconstructed GST. At the global scale, spatial and temporal sampling aggregated effect leads to underestimating the simulated GST 20th-century increase. Most of this reduced skill in capturing the 20th-century warming is the result of temporal logging of the borehole records because a large portion of the BTPs do not contain information about the warming during the last few decades of the 20th century as they are dated before 1982. In the CESM-LME simulations, this effect can range between > 0 and 0.3 °C century^{-1}, representing up to 20% of the total warming.

In the surrogate reality of the CESM-LME, the superposition of the physical SAT–GST decoupling and the methodological issues further enhance underestimating the 20th century SAT increase as represented by the pseudo-reconstructed GST. The SAT estimated by the reconstructed GST, can be additionally impacted by the physical processes that interrupt the long-term SAT-GST coupling. The PPE implementation shows that even in the idealized scenario, in which sampling is not limited in space or time, the global SAT warming during industrial times is not fully captured by the GST pseudo-reconstruction with differences than can reach around 20% of the simulated warming on average in the ALL-F ensemble. Globally, the GHG external forcing is the dominant effect contributing to the larger SAT increase relative to GST during industrial times. At smaller spatial scales, other external forcings take over the SAT–GST decoupling. For instance, the LULC forcing represents the dominant effect over the European and the African regions. The aggregated effect of the long-term SAT–GST decoupling and the one imposed by the methodological aspects further enhance the potential underestimation of the 20th century SAT. On a global scale, such a combined effect can miss up to 50% of the simulated SAT (average values for the ensemble)

with the largest contributions from the methodological aspects.

Internal variability within the CESM-LME world plays an important role in the level of impact of both methodological and physical biases on borehole-based reconstructions. When considering the ensemble that incorporates all natural and anthropogenic forcings, the GST (SAT) warming varies among ensemble members ranging between 0.28 and 0.51 (0.35 and 0.60) °C century^{-1}. This inter-simulation variability reflects the internal variability of the system as changing the initial conditions to generate the ensemble results in a different trajectory of the ensemble members. This range of variability is also found regarding the differences between simulated and pseudo-reconstructed results. While some simulations experience no significant change, in others, masked PPE can experience underestimations of several tenths of a degree depending on the realization of internal variability.

In light of these findings, both the methodological issues and the borehole-based reconstructions physical biases would lead to an underestimation of the temperature increase from the LIA to the present day, especially over the industrial period. These results cannot attribute the larger temperature increase suggested by the actual borehole estimations during the last few centuries of the LM relative to those of some other proxy-based reconstructions to methodological biases.

6.2 *Outlook*

The development of SoTOS offers a unique opportunity to evaluate the performance of climate model simulations on representing the temporal and spatial variability of soil temperatures. This is relevant because soil temperature is a key variable that influences the representation of many physical processes in the atmosphere and below the surface in numerical weather prediction, regional modeling and land-reanalysis approaches. Therefore, it is worth that future work focuses on evaluating the accuracy of numerical simulations on reproducing this variable and how this impact the simulation of the surface energy balance. Performing regional-climate model simulations with different parameterizations can provide insights into the main mechanisms that influence the representation of soil temperature. For instance, the development of high-resolution model configurations over Spain can help assessing the response to local processes that limit the representation of soil temperature from relatively lower resolution configurations, as shown in this work. Similarly, testing different land-surface model configurations,

for instance, deeper BBCPs, can enhance our understanding of the influence of the BBCP on representing the propagating temperature signals with depth.

This work has shown that the combined effect of the methodological constraints and the physical SAT–GST decoupling leads to a systematic underestimation of the 20th-century SAT trends at a global scale. However, even if the analysis included herein represents the most realistic implementation to date of the borehole method in PPEs, further analyses are desirable to enhance our understanding of other factors that may influence the level of impact. For instance, the use of other ESMs with different climate sensitivities, i.e., larger 20th-century warming, and a different representation of the influence of changes in land surface physics (e.g., snow cover or soil moisture) could lead to a different representation of changes in the SAT–GST relationship. Additionally, the limitations in the local representation of sampling due to model resolution and more technical issues like the existence of local noise in BTPs have not been considered here. Therefore, exploring these issues in future works will allow for a more complete evaluation of the method. The latter would be especially interesting if new ensembles of LM simulations with ESMs including both all- and single-forcing experiments were developed in the frame of the CMIP6/PMIP4 (Coupled Model Intercomparison Project phase 6/Paleoclimate Modeling Intercomparison Project phase 4; Eyring et al., 2016; Jungclaus et al., 2017, respectively). This would allow an exploration of the influence of different external forcing factors and different model physics that have some influence on, for instance, SAT–GST decoupling. To date, this issue can only be addressed using the CESM-LME, as we have done in this Thesis. Additionally, this work clearly supports the need for updating and expanding the borehole network. More and, if possible, deeper and good-quality BTPs are needed.

Finally, this work has evidenced the large influence of internal variability when analyzing the performance of borehole-based reconstructions on retrieving the 20th-century trends. Addressing this effect in multi-model ensembles can improve our understanding of the internal variability influence on the overall trajectory of SAT and GST during the 20th century, considering the actual BTPs spatial and temporal distribution. Furthermore, the use of data from long-term reanalyses covering the full 20th century (e.g., Poli et al., 2016; Laloyaux et al., 2016) or even extending back to the mid 19th century (Slivinski et al., 2019) also arises as a valuable opportunity to evaluate the role of internal variability on capturing the SAT evolution from reconstructed GST considering the distribution of BTPs at global or regional scales.

References

Alexeev, V. A., D. J. Nicolsky, V. E. Romanovsky, and D. M. Lawrence, 2007: An evaluation of deep soil configurations in the clm3 for improved representation of permafrost. *Geophysical Research Letters*, **34**, doi:10.1029/2007GL029536, l09502.

Anderson, R., J. Canadell, J. Randerson, R. Jackson, B. Hungate, D. Baldocchi, G. Ban-Weiss, G. Bonan, K. Caldeira, L. Cao, N. Diffenbaugh, K. Gurney, L. Kueppers, B. Law, S. Luyssaert, and T. O'Halloran, 2011: Biophysical considerations in forestry for climate protection. *Frontiers in Ecology and the Environment*, **9**, 174–182, doi:10.1890/090179.

Baker, J. M. and D. G. Baker, 2002: Long–term ground heat flux and heat storage at a mid-latitude site. *Climatic Change*, **54**, 295–303, doi:10.1023/A:1016144718218.

Balsamo, G., C. Albergel, A. Beljaars, S. Boussetta, E. Brun, H. Cloke, D. Dee, E. Dutra, J. Muñoz Sabater, F. Pappenberger, P. de Rosnay, T. Stockdale, and F. Vitart, 2015: Era-interim/land: a global land surface reanalysis data set. *Hydrology and Earth System Sciences*, **19**, 389–407, doi:10.5194/hess-19-389-2015.

Bartlett, M. G., D. S. Chapman, and R. N. Harris, 2004: Snow and the ground temperature record of climate change. *J. Geophys. Res.*, **109**, F04008, doi:10.1029/2004JF000224.

— 2005: Snow effect on north american ground temperatures, 1950-2002. *J. Geophys. Res.*, **110**, F03008, doi:10.1029/2005JF000293.

Beck, A. E., P. Y. Shen, H. Beltrami, J. C. Mareschal, J. Safanda, M. N. Sebagenzi, G. Vasseur, and K. Wang, 1992: A comparison of five different analyses in the interpretation of five borehole temperature data sets.

Palaeogeography, Palaeoclimatology, Palaeoecology (GPC section), **98**, 101–112, doi:10.1016/j.epsl.2004.09.014.

Bell, J. E., M. A. Palecki, C. B. Baker, W. G. Collins, J. H. Lawrimore, R. D. Leeper, M. E. Hall, J. Kochendorfer, T. P. Meyers, T. Wilson, and H. J. Diamond, 2013: U.S. Climate Reference Network Soil Moisture and Temperature Observations. *Journal of Hydrometeorology*, **14**, 977–988, doi:10.1175/JHM-D-12-0146.1.

Beltrami, H. and E. Bourlon, 2004: Ground warming patterns in the northern hmeisphere during the last five centuries. *Earth Planet. Sci. Lett.*, **227**, 169–177, doi:10.1016/j.epsl.2004.09.014.

Beltrami, H. and L. Kellman, 2003: An examination of short- and long-term airground temperature coupling. *Global and Planetary Change*, **38**, 291 – 303, doi:10.1016/S0921-8181(03)00112-7.

Beltrami, H. and J. C. Mareschal, 1991: Recent warming in eastern canada inferred from geothermal measurements. *Geophys. Res. Lett.*, **18(4)**, 605–608.

— 1995: Resolution of ground temperature histories inverted from borehole temperature data. *Glob. Planet. Change*, **11**, 57–70.

Beltrami, H., G. S. Matharoo, and J. E. Smerdon, 2015: Impact of borehole depths on reconstructed estimates of ground surface temperature histories and energy storage. *Journal of Geophysical Research: Earth Surface*, **120**, 763–778, doi:10.1002/2014JF003382.

Beltrami, H., J. E. Smerdon, G. S. Matharoo, and N. Nickerson, 2011: Impact of maximum borehole depths on inverted temperature histories in borehole paleoclimatology. *Climate of the Past*, **7**, 745–756, doi:10.5194/cp-7-745-2011.

Berger, A., M.-F. Loutre, and C. Tricot, 1993: Insolation and earth's orbital periods. *Journal of Geophysical Research: Atmospheres*, **98**, 10341–10362, doi:10.1029/93JD00222.

Bodri, L. and V. Cermak, 2005: Borehole temperatures, climate change and the pre-observational surface air temperature mean: allowance for hydraulic conditions. *Global and Planetary Change*, **45**, 265 – 276, doi:10.1016/j.gloplacha.2004.09.001.

— 2007a: *Background and History of the Problem. In Borehole Climatology: a new method how to reconstruct climate*, Elsevier Science Ltd, Oxford. 1–36, doi:10.1016/B978-008045320-0/50001-6.

— 2007b: *Climate Change and Subsurface Temperature. In Borehole Climatology: a new method how to reconstruct climate*, Elsevier Science Ltd, Oxford. 37–173, doi:10.1016/B978-008045320-0/50002-8.

Bonan, G., 2015: *Surface Energy Fluxes. In Ecological Climatology: Concepts and Applications*, Cambridge University Press. 3 edition, 193–208, doi:10.1017/CBO9781107339200.013.

Bond-Lamberty, B. and A. Thomson, 2010: Temperature-associated increases in the global soil respiration record. *Nature*, **464**, 579–582, doi:10.1038/nature08930.

Brovkin, V., M. Claussen, E. Driesschaert, T. Fichefet, D. Kicklighter, M. F. Loutre, H. D. Matthews, N. Ramankutty, M. Schaeffer, and A. Sokolov, 2006: Biogeophysical effects of historical land cover changes simulated by six earth system models of intermediate complexity. *Climate Dynamics*, **26**, 587–600, doi:10.1007/s00382-005-0092-6.

Carey, J. C., J. Tang, P. H. Templer, K. D. Kroeger, T. W. Crowther, A. J. Burton, J. S. Dukes, B. Emmett, S. D. Frey, M. A. Heskel, L. Jiang, M. B. Machmuller, J. Mohan, A. M. Panetta, P. B. Reich, S. Reinsch, X. Wang, S. D. Allison, C. Bamminger, S. Bridgham, S. L. Collins, G. de Dato, W. C. Eddy, B. J. Enquist, M. Estiarte, J. Harte, A. Henderson, B. R. Johnson, K. S. Larsen, Y. Luo, S. Marhan, J. M. Melillo, J. Peñuelas, L. Pfeifer-Meister, C. Poll, E. Rastetter, A. B. Reinmann, L. L. Reynolds, I. K. Schmidt, G. R. Shaver, A. L. Strong, V. Suseela, and A. Tietema, 2016: Temperature response of soil respiration largely unaltered with experimental warming. *Proceedings of the National Academy of Sciences*, **113**, 13797–13802, doi:10.1073/pnas.1605365113.

Carslaw, H. S. and J. C. Jaeger, 1959: *Conduction of heat in solids*. Oxford Univ. Press. 2nd ed., New York, 3rd. edition.

Cermak, V., 1971: Underground temperature and inferred climatic temperature of the past millenium. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **10**, 1 – 19, doi:10.1016/0031-0182(71)90043-5.

Cermak, V. and L. Bodri, 2018: Attribution of precipitation changes on ground–air temperature offset: Granger causality analysis. *International Journal of Earth Sciences*, **107**, 145–152, doi:10.1007/s00531-016-1351-y.

Cermak, V., L. Bodri, M. Kresl, P. Dedecek, and J. Safanda, 2017: Eleven years of groundair temperature tracking over different land cover types. *International Journal of Climatology*, **37**, 1084–1099, doi:10.1002/joc.4764.

Chisholm, T. J. and D. S. Chapman, 1992: Climate change inferred from analysis of borehole temperatures: An example from western utah. *Journal of Geophysical Research: Solid Earth*, **97**, 14155–14175, doi:10.1029/92JB00765.

Claussen, M., V. Brovkin, and A. Ganopolski, 2001: Biogeophysical versus biogeochemical feedbacks of large-scale land cover change. *Geophysical Research Letters*, **28**, 1011–1014, doi:10.1029/2000GL012471.

Cooper, P. J. M. and R. Law, 1978: Enhanced soil temperature during very early growth and its association with maize development and yield in the highlands of kenya. *The Journal of Agricultural Science*, **91**, 569577, doi:10.1017/S0021859600059955.

Cosenza, P., R. Guérin, and A. Tabbagh, 2003: Relationship between thermal conductivity and water content of soils using numerical modelling. *European Journal of Soil Science*, **54**, 581–588, doi:10.1046/j.1365-2389.2003.00539.x.

Cuesta-Valero, F. J., A. Garca-Garca, H. Beltrami, and J. E. Smerdon, 2016: First assessment of continental energy storage in cmip5 simulations. *Geophysical Research Letters*, **43**, 5326–5335, doi:10.1002/2016GL068496.

Davidson, E. A. and I. A. Janssens, 2006: Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature*, **440**, 165–173, doi:10.1038/nature04514.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. *Geoscientific Model Development*, **9**, 1937–1958, doi:10.5194/gmd-9-1937-2016.

Farouki, O. T., 1986: *Thermal Properties of Soils. Series on Rock and Soil Mechanics*. Trans Tech Publications Limited, United States, 3rd. edition.

Ferguson, G., H. Beltrami, and A. D. Woodbury, 2006: Perturbation of ground surface temperature reconstructions by groundwater flow? *Geophys. Res. Lett.*, **33**, L13708, doi:10.1029/2006GL026634.

Fernández-Donado, L., J. F. González-Rouco, C. C. Raible, C. M. Ammann, D. Barriopedro, E. García-Bustamante, J. H. Jungclaus, S. J. Lorenz, J. Luterbacher, S. J. Phipps, J. Servonnat, D. Swingedouw, S. F. B. Tett, S. Wagner, P. Yiou, and E. Zorita, 2013: Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Climate of the Past*, **9**, 393–421, doi:10.5194/cp-9-393-2013.

Gallego-Elvira, B., C. M. Taylor, P. P. Harris, D. Ghent, K. L. Veal, and S. S. Folwell, 2016: Global observational diagnosis of soil moisture control on the land surface energy balance. *Geophysical Research Letters*, **43**, 2623–2631, doi:10.1002/2016GL068178.

Gao, C. C., A. Robock, and C. Ammann, 2008a: Volcanic forcing of climate over the past 1500 years: an improved ice core-based index for climate models. *J. Geophys. Res.*, **113**, D23111, doi: 10.1029/2008JD010239.

Gao, Z., D. H. Lenschow, R. Horton, M. Zhou, L. Wang, and J. Wen, 2008b: Comparison of two soil temperature algorithms for a bare ground site on the loess plateau in china. *Journal of Geophysical Research: Atmospheres*, **113**, doi:https://doi.org/10.1029/2008JD010285.

García-García, A., F. J. Cuesta-Valero, H. Beltrami, and J. E. Smerdon, 2016: Simulation of air and ground temperatures in pmip3/cmip5 last millennium simulations: implications for climate reconstructions from borehole temperature profiles. *Environmental Research Letters*, **11**, 044022, doi:10.1088/1748-9326/11/4/044022.

García-García, A., F. J. Cuesta-Valero, H. Beltrami, and J. E. Smerdon, 2019: Characterization of air and ground temperature relationships within the cmip5 historical and future climate simulations. *Journal of Geophysical Research: Atmospheres*, **124**, 3903–3929, doi:10.1029/2018JD030117.

Geiger, R., R. H. Aron, and T. Paul, 1995: *Earths Surface Energy Budget. In: The Climate Near the Ground*, Vieweg+Teubner Verlag. 5th edition.

Georgescu, M., G. Miguez-Macho, L. Steyaert, and C. Weaver, 2008: Sensitivity of summer climate to anthropogenic land-cover change over the greater phoenix, az, region. *Journal of Arid Environments*, **72**, 1358 – 1373.

Gonzalez-Hidalgo, J. C., D. Peña Angulo, M. Brunetti, and N. Cortesi, 2015: Motedas: a new monthly temperature database for mainland spain and the trend in temperature (1951–2010). *International Journal of Climatology*, **35**, 4444–4463, doi:10.1002/joc.4298.

González-Rouco, J. F., H. Beltrami, E. Zorita, and B. Stevens, 2009: Borehole climatology: a discussion based on contributions from climate modelling. *Clim. Past*, **5**, 99–127, doi:10.5194/cp-5-97-2009.

González-Rouco, J. F., H. Beltrami, E. Zorita, and H. von Storch, 2006: Simulation and inversion of borehole temperature profiles in surrogate climates: Spatial distribution and surface coupling. *Geophys. Res. Lett.*, **33**, L01703, doi:10.1029/2005GL024693.

González-Rouco, J. F., J. L. Jiménez, V. Quesada, and F. Valero, 2001: Quality control and homogeneity of precipitation data in the southwest of europe. *J. Clim.*, **14**, 964–978, doi:10.1175/1520-0442(2001)014<0964:QCAHOP>2.0.CO;2.

González-Rouco, J. F., H. von Storch, and E. Zorita, 2003: Deep soil temperature as proxy for surface air-temperature in a coupled model simulation of the last thousand years. *Geophysical Research Letters*, **30**, doi:10.1029/2003GL018264.

Hao, G., Q. Zhuang, J. Pan, Z. Jin, X. Zhu, and S. Liu, 2014: Soil thermal dynamics of terrestrial ecosystems of the conterminous united states from 1948 to 2008: an analysis with a process-based soil physical model and ameriflux data. *Climatic Change*, **126**, 135–150, doi:10.1007/s10584-014-1196-y.

Harris, R. N. and D. S. Chapman, 1995: Climate change on the colorado plateau of eastern utah inferred from borehole temperatures. *Journal of Geophysical Research: Solid Earth*, **100**, 6367–6381, doi:10.1029/94JB02165.

— 2001: Mid-latitude (30-60 N) climatic warming inferred by combining borehole temperatures with surface air temperatures. *Geophysical Research Letters*, **28**, 747–750, doi:10.1029/2000GL012348.

Hartmann, D., A. KleinTank, M. Rusticucci, L. Alexander, S. Brönnimann, Y. Charabi, F. Dentener, E. Dlugokencky, D. Easterling, A. Kaplan, B. Soden, P. Thorne, M. Wild, and P. Zhai, 2013: *Observations: Atmosphere and*

Surface, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, book section 2. 159254, doi:10.1017/CBO9781107415324.008.

Hegerl, G., F. Zwiers, and C. Tebaldi, 2011: Patterns of change: whose fingerprint is seen in global warming? *Environmental Research Letters*, **6**, –, doi:10.1088/1748-9326/6/4/044025.

Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Hornyi, J. Muoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. De Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hlm, M. Janiskov, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume, and J.-N. Thpaut, 2020: The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146**, 1999–2049, doi:10.1002/qj.3803.

Hinkel, K., F. Paetzold, F. Nelson, and J. Bockheim, 2001: Patterns of soil temperature and moisture in the active layer and upper permafrost at barrow, alaska: 19931999. *Global and Planetary Change*, **29**, 293 – 309, doi:10.1016/S0921-8181(01)00096-0.

Hu, Q. and S. Feng, 2003: A daily soil temperature dataset and soil temperature climatology of the contiguous united states. *Journal of Applied Meteorology*, **42**, 1139–1156, doi:10.1175/1520-0450(2003)042<1139:ADSTDA>2.0.CO;2.

Huang, R., J. xi Huan, C. Zhang, H. yuan Ma, W. Zhuo, Y. yi Chen, D. hai Zhu, Q. Wu, and L. R. Mansaray, 2020: Soil temperature estimation at different depths, using remotely-sensed data. *Journal of Integrative Agriculture*, **19**, 277 – 290, doi:10.1016/S2095-3119(19)62657-2.

Huang, S. and H. N. Pollack, 1998: Global borehole temperature database for climate reconstruction. *IGBP PAGES/World Data Center-A for Paleoclimatology Data Contribution Series 1998-044. NOAA/NGDC Paleoclimatology Program, Boulder CO, USA. GR1*.

Huang, S., H. N. Pollack, and P. Y. Shen, 2000: Temperature trends over the past five centuries reconstructed from borehole temperatures. *Nature*, **403**, 756–758, doi:10.1038/35001556.

Huang, S., P. Y. Shen, and H. N. Pollack, 1996: Deriving century-long trends of surface temperature change from borehole temperatures. *Geophysical Research Letters*, **23**, 257–260, doi:10.1029/96GL00020.

Hunke, E. C., W. H. Lipscomb, A. K. Turner, N. Jeffery, and S. Elliott, 2015: Cice: the los alamos sea ice model documentation and software users manual version 5.1. (Accessed 2020-04-03).

Hurrell, J. W., M. M. Holland, P. R. Gent, S. Ghan, J. E. Kay, P. J. Kushner, J.-F. Lamarque, W. G. Large, D. Lawrence, K. Lindsay, W. H. Lipscomb,

M. C. Long, N. Mahowald, D. R. Marsh, R. B. Neale, P. Rasch, S. Vavrus, M. Vertenstein, D. Bader, W. D. Collins, J. J. Hack, J. Kiehl, and S. Marshall, 2013: The community earth system model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, **94**, 1339–1360, doi:10.1175/BAMS-D-12-00121.1.

Hurtt, G. C., L. P. Chini, S. Frolking, R. Betts, J. Feedema, G. Fischer, K. K. Goldewijk, K. Hibbard, A. Janetos, C. Jones, G. Kindermann, T. Kinoshita, K. Riahi, E. S. andS. Smith, E. Stehfest, A. Thomson, P. Thorton, D. van Vuuren, and Y. Wang, 2009: Harmonization of global land use scenarios for the period 1500-2100 for ipcc-ar5. *Integrated Land Ecosystem-Atmosphere Processes Study (iLEAPS) Newsletter*, **7**, 6–8.

Hurtt, G. C., L. P. Chini, R. Sahajpal, S. E. Frolking, J. Fisk, B. Bodirsky, K. V. Calvin, S. Fujimori, K. Goldewijk, T. Hasegawa, P. Havlik, A. Heinimann, F. Humpenöder, J. O. Kaplan, T. Krisztin, D. M. Lawrence, P. Lawrence, O. Mertz, A. Popp, K. Riahi, E. Stehfest, D. van Vuuren, L. de Waal, and X. Zhang, 2016: Harmonization of global land-use scenarios for the period 850-2100. *AGU Fall Meeting Abstracts*, volume 2016, GC23I–01.

IFS, 2018: *ECMWF Integrated Forecast System Documentation. Part IV : Physical processes*, ECMWF. Number 4 in IFS Documentation, https://www.ecmwf.int/node/18714 (Accessed: 2020-08-30).

Jackson, R. B., J. T. Randerson, J. G. Canadell, R. G. Anderson, R. Avissar, D. D. Baldocchi, G. B. Bonan, K. Caldeira, N. S. Diffenbaugh, C. B. Field, B. A. Hungate, E. G. Jobbgy, L. M. Kueppers, M. D. Nosetto, and D. E. Pataki, 2008: Protecting climate with forests. *Environmental Research Letters*, **3**, 044006, doi:10.1088/1748-9326/3/4/044006.

Jansen, E., J. Overpeck, K. R. Briffa, J. C. Duplessy, V. Masson-Delmontte, D. Olago, B. Otto-Bliesner, W. R. Peltier, S. Rahmstorf, R. Ramesh, D. Raynaud, D. Rind, O. Solomina, R. Villalba, and D. Zhang, 2007: *Paleoclimate. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [S. Solomon and D. Qin and M. Manning and Z. Chen and M. Marquis and K. B. Averyt and M. Tignor and H. L. Miller (eds.)]*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jaume-Santero, F., C. Pickler, H. Beltrami, and J.-C. Mareschal, 2016: North american regional climate reconstruction from ground surface temperature histories. *Climate of the Past*, **12**, 2181–2194, doi:10.5194/cp-12-2181-2016.

Jones, P. D., K. R. Briffa, T. J. Osborn, J. M. Lough, T. D. Van Ommen, B. M. Vinther, J. Luterbacher, E. R. Wahl, F. W. Zwiers, M. E. Mann, G. A. Schmidt, C. M. Ammann, B. M. Buckley, K. M. Cobb, J. Esper, H. Goose, N. Graham,

E. Jansen, T. Kiefer, C. Kull, M. Kttel, E. Mosley-Thompson, J. T. Overpeck, N. Riedwyl, M. Schulz, A. W. Tudhope, R. Villalba, H. Wanner, E. Wilff, and E. Xoplake, 2009: High-resolution palaeoclimatology of the last millennium: A review of current status and future prospects. *Holocene*, **19**, 3–49, doi:10.1177/0959683608098952.

Jungclaus, J. H., E. Bard, M. Baroni, P. Braconnot, J. Cao, L. P. Chini, T. Egorova, M. Evans, J. F. González-Rouco, H. Goosse, G. C. Hurtt, F. Joos, J. O. Kaplan, M. Khodri, K. Klein Goldewijk, N. Krivova, A. N. LeGrande, S. J. Lorenz, J. Luterbacher, W. Man, A. C. Maycock, M. Meinshausen, A. Moberg, R. Muscheler, C. Nehrbass-Ahles, B. I. Otto-Bliesner, S. J. Phipps, J. Pongratz, E. Rozanov, G. A. Schmidt, H. Schmidt, W. Schmutz, A. Schurer, A. I. Shapiro, M. Sigl, J. E. Smerdon, S. K. Solanki, C. Timmreck, M. Toohey, I. G. Usoskin, S. Wagner, C.-J. Wu, K. L. Yeo, D. Zanchettin, Q. Zhang, and E. Zorita, 2017: The pmip4 contribution to cmip6 – part 3: The last millennium, scientific objective, and experimental design for the pmip4 past1000 simulations. *Geoscientific Model Development*, **10**, 4005–4033, doi:10.5194/gmd-10-4005-2017.

Jungqvist, G., S. Oni, C. Teutschbein, and M. Futter, 2014: Effect of climate change on soil temperature in swedish boreal forests. *PLoS ONE*, **9**, doi:10.1371/journal.pone.0093957.

Kane., D., K. Hinkel., D. Goering., L. Hinzman., and S. Outcalt, 2001: Nonconductive heat transfer associated with frozen soils. *Global and Planetary Change*, **29**, 275–292.

Kaspar, T. C. and W. L. Bland, 1992: Soil temperature and root growth. *Soil Science*, **154**, doi:10.1097/00010694-199210000-00005.

Killick, R., I. A. Eckley, K. Ewans, and P. Jonathan, 2010: Detection of changes in variance of oceanographic time-series using changepoint analysis. *Ocean Engineering*, **37**, 1120 – 1126, doi:10.1016/j.oceaneng.2010.04.009.

Kluitenberg, G., 2002: 5.2 heat capacity and specific heat. *Methods of Soil Analysis: Part 4 Physical Methods*, **5**, 1201–1208.

Kohl, T., 1998: Paleoclimatic temperature signals — can they be washed out? *Tectonophysics*, **291**, 225 – 234, doi:10.1016/S0040-1951(98)00042-0.

Kukkonen, I. T. and C. Clauser, 1994: Simulation of heat transfer at the kola deep-hole site- implications for advection, heat refraction and paleoclimatic effects. *Geophys. J. Int.*, **41**, 175–191, doi:10.1111/j.1365-246X.1994.tb01806.x.

Lachenbruch, A. H. and B. V. Marshall, 1986: Changing climate: geothermal evidence from permafrost in the alaskan arctic. *Science*, **234**, 689–696, doi:10.1126/science.234.4777.689.

Laloyaux, P., M. Balmaseda, D. Dee, K. Mogensen, and P. Janssen, 2016: A coupled data assimilation system for climate reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **142**, 65–78, doi:10.1002/qj.2629.

Landrum, L., B. L. Otto-Bliesner, E. R. Wahl, A. Conley, P. J. Lawrence, N. Rosenbloom, and H. Teng, 2013: Last millennium climate and its variability in ccsm4. *Journal of Climate*, **26**, 1085–1111, doi:10.1175/JCLI-D-11-00326.1.

Lane, E. C., 1923: Geotherms of the lake superior copper country. *Bull. Geol. Soc. Am.*, **34**, 703–720, doi:10.1130/GSAB-34-703.

Lawrence, D. M., K. W. Oleson, M. G. Flanner, P. E. Thornton, S. C. Swenson, P. J. Lawrence, X. Zeng, Z.-L. Yang, S. Levis, K. Sakaguchi, G. B. Bonan, and A. G. Slater, 2011: Parameterization improvements and functional and structural advances in version 4 of the community land model. *Journal of Advances in Modeling Earth Systems*, **3**, doi:10.1029/2011MS00045, m03001.

Lee, X., M. L. Goulden, D. Y. Hollinger, A. Barr, T. A. Black, G. Bohrer, R. Bracho, B. Drake, A. Goldstein, L. Gu, G. Katul, T. Kolb, B. E. Law, H. Margolis, T. Meyers, R. Monson, W. Munger, R. Oren, K. T. Paw U, A. D. Richardson, H. P. Schmid, R. Staebler, S. Wofsy, and L. Zhao, 2011: Observed increase in local cooling effect of deforestation at higher latitudes. *Nature*, **479**, 384–387, doi:10.1038/nature10588.

Legutke, S. and R. Voss, 1999: The hamburg atmosphere-ocean coupled circulation model echo-g. Technical Report 18, DKRZ, Hamburg.

Lewis, T. and K. Wang, 1992: Influence of terrain on bedrock temperatures. *Global and Planetary Change*, **6**, 87 – 100, doi:10.1016/0921-8181(92)90028-9.

Li, M., P. Wu, and Z. Ma, 2020: A comprehensive evaluation of soil moisture and soil temperature from third-generation atmospheric and land reanalysis data sets. *International Journal of Climatology*, 1–23, doi:10.1002/joc.6549.

Lucio-Eceiza, E. E., J. F. Gonzlez-Rouco, J. Navarro, and H. Beltrami, 2018a: Quality Control of Surface Wind Observations in Northeastern North America. Part I: Data Management Issues. *Journal of Atmospheric and Oceanic Technology*, **35**, 163–182, doi:10.1175/JTECH-D-16-0204.1.

Lucio-Eceiza, E. E., J. F. Gonzlez-Rouco, J. Navarro, H. Beltrami, and J. Conte, 2018b: Quality Control of Surface Wind Observations in Northeastern North America. Part II: Measurement Errors. *Journal of Atmospheric and Oceanic Technology*, **35**, 183–205, doi:10.1175/JTECH-D-16-0205.1.

MacDougall, A. H. and H. Beltrami, 2017: Impact of deforestation on subsurface temperature profiles: implications for the borehole paleoclimate record. *Environmental Research Letters*, **12**, 074014, doi:10.1088/1748-9326/aa7394.

MacFarling Meure, C., D. Etheridge, C. Trudinger, P. Steele, R. Langenfelds, T. Van Ommen, A. Smith, and J. Elkins, 2006: Law Dome CO2, CH4 and N2O ice core records extended to 2000 years BP. *Geophys. Res. Lett.*, **33**, L14810, doi:10.1029/2006GL026152.

Mankin, J. S., R. Seager, J. E. Smerdon, B. I. Cook, A. P. Williams, and R. M. Horton, 2018: Blue water trade-offs with vegetation in a co2-enriched climate.

Geophysical Research Letters, **45**, 3115–3125, doi:10.1002/2018GL077051.

Mann, M. E. and G. A. Schmidt, 2003: Ground vs. surface air temperature trends: implications for borehole surface temperature reconstructions. *Geophys. Res. Let.*, **30**, 1607, doi:10.1029/2003GL017170.

Mareschal, J.-C. and H. Beltrami, 1992: Evidence for recent warming from perturbed geothermal gradients: examples from eastern canada. *Climate Dynamics*, **6**, 135–143, doi:10.1007/BF00193525.

Masson-Delmotte, V., M. Schulz, A. Abe-Ouchi, J. Beer, A. Ganopolski, J. GonzálezRouco, E. Jansen, K. Lambeck, J. Luterbacher, T. Naish, T. Osborn, B. Otto-Bliesner, T. Quinn, R. Ramesh, M. Rojas, X. Shao, and A. Timmermann, 2013: *Information from Paleoclimate Archives*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, book section 5. 383464.

McCumber, M. C. and R. A. Pielke, 1981: Simulation of the effects of surface fluxes of heat and moisture in a mesoscale numerical model: 1. soil layer. *Journal of Geophysical Research: Oceans*, **86**, 9929–9938, doi:10.1029/JC086iC10p09929.

McMichael, B. and J. Burke, 1998: Soil temperature and root growth. *HortScience HortSci*, **33**, 947 – 951.

Meek, D. and J. Hatfield, 1994: Data quality checking for single station meteorological databases. *Agricultural and Forest Meteorology*, **69**, 85 – 109, doi:10.1016/0168-1923(94)90083-3, new Visions and New Perspectives in Agricultural Meteorology/Climatology and Modeling.

Melillo, J., J. Steudler, P. Aber, K. Newkirk, H. Lux, F. Bowles, C. Catricala, A. Magill, T. Ahrens, and S. Morrisseau, 2002: Soil warming and carbon-cycle feedbacks to the climate system. *Science*, **298**, 2173–6, doi:10.1126/science.1074153.

Melo-Aguilar, C., J. F. González-Rouco, E. García-Bustamante, J. Navarro-Montesinos, and N. Steinert, 2018: Influence of radiative forcing factors on ground–air temperature coupling during the last millennium: implications for borehole climatology. *Climate of the Past*, **14**, 1583–1606, doi:10.5194/cp-14-1583-2018.

Melo-Aguilar, C., J. F. González-Rouco, E. García-Bustamante, N. Steinert, J. H. Jungclaus, J. Navarro, and P. J. Roldán-Gómez, 2020: Methodological and physical biases in global to subcontinental borehole temperature reconstructions: an assessment from a pseudo-proxy perspective. *Climate of the Past*, **16**, 453–474, doi:10.5194/cp-16-453-2020.

Muñoz Sabater, J., 2019: Era5-land monthly averaged data from 1981 to present. copernicus climate change service (c3s) climate data store. 10.24381/cds.68d2bb30, accessed: 2020-09-16.

Myhre, G., D. Shindell, F.-M. Bréon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, T. Takemura, and H. Zhang, 2013: *Anthropogenic and Natural Radiative Forcing*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, book section 8. 659740.

Neale, R. B., A. Gettelman, S. Park, A. J. Conley, D. Kinnison, D. Marsh, A. K. Smith, F. Vitt, H. Morrison, P. Cameron-smith, W. D. Collins, M. J. Iacono, R. C. Easter, X. Liu, M. A. Taylor, C. chieh Chen, P. H. Lauritzen, D. L. Williamson, R. Garcia, J. francois Lamarque, M. Mills, S. Tilmes, S. J. Ghan, P. J. Rasch, and M. Meteorology, 2012: Description of the ncar community atmosphere model (cam 5.0), tech. note ncar/tn-486+str, natl. cent. for atmos. 1–289.

Neukom, R., L. A. Barboza, M. P. Erb, F. Shi, J. Emile-Geay, M. N. Evans, J. Franke, D. S. Kaufman, L. Lcke, K. Rehfeld, A. Schurer, F. Zhu, S. Brnnimann, G. J. Hakim, B. J. Henley, F. C. Ljungqvist, N. McKay, V. Valler, L. von Gunten, and P. 2k Consortium, 2019: Consistent multidecadal variability in global temperature reconstructions and simulations over the common era. *Nature Geoscience*, **12**, 643–649, doi:10.1038/s41561-019-0400-0.

Nicolsky, D. J., V. E. Romanovsky, V. A. Alexeev, and D. M. Lawrence, 2007: Improved modeling of permafrost dynamics in a gcm land-surface scheme. *Geophys. Res. Lett.*, **34**, L8501,doi:10.1029/2007GL029525.

Ochsner, T. E., R. Horton, and T. Ren, 2001: A new perspective on soil thermal properties. *Soil Science Society of America Journal*, **65**, 1641–1647, doi:10.2136/sssaj2001.1641.

Oladunjoye, M. A. and O. A. Sanuade, 2012: Thermal diffusivity, thermal effusivity and specific heat of soils in olorunsogo powerplant, southwestern nigeria. volume 13, 502–521.

Oleson, K. W., D. M. Lawrence, G. B, M. G. Flanner, E. Kluzek, P. J, S. Levis, S. C. Swenson, E. Thornton, J. Feddema, C. L. Heald, J. francois Lamarque, G. yue Niu, T. Qian, S. Running, K. Sakaguchi, L. Yang, X. Zeng, X. Zeng, and M. Decker, 2010: Technical description of version 4.0 of the community land model (clm).

Oleson, K. W., G.-Y. Niu, Z.-L. Yang, D. M. Lawrence, P. E. Thornton, P. J. Lawrence, R. Stckli, R. E. Dickinson, G. B. Bonan, S. Levis, A. Dai, and T. Qian, 2008: Improvements to the community land model and their impact on the hydrological cycle. *Journal of Geophysical Research: Biogeosciences*, **113**, G01021, doi:10.1029/2007JG000563.

Otto-Bliesner, B. L., E. C. Brady, J. Fasullo, A. Jahn, L. Landrum, S. Stevenson, N. Rosenbloom, A. Mai, and G. Strand, 2016: Climate variability and change since 850 ce: An ensemble approach with the community earth sys-

tem model. *Bulletin of the American Meteorological Society*, **97**, 735–754, doi:10.1175/BAMS-D-14-00233.1.

Outcalt, S. I., F. E. Nelson, and K. M. Hinkel, 1990: The zero-curtain effect: Heat and mass transfer across an isothermal region in freezing soil. *Water Resources Research*, **26**, 1509–1516, doi:10.1029/WR026i007p01509.

Page, E. S., 1955: A test for a change in a parameter occurring at an unknown point. *Biometrika*, **42**, 523–527, doi:10.1093/biomet/42.3-4.523.

Patil, M., R. Waghmare, S. Halder, and T. Dharmaraj, 2011: Performance of noah land surface model over the tropical semi-arid conditions in western india. *Atmospheric Research*, **99**, 85 – 96, doi:10.1016/j.atmosres.2010.09.006.

Peterson, T. C., R. Vose, R. Schmoyer, and V. Razuvav, 1998: Global historical climatology network (ghcn) quality control of monthly temperature data. *International Journal of Climatology*, **18**, 1169–1179, doi:10.1002/(SICI)1097-0088(199809)18:11<1169::AID-JOC309>3.0.CO;2-U.

Poli, P., H. Hersbach, D. P. Dee, P. Berrisford, A. J. Simmons, F. Vitart, P. Laloyaux, D. G. H. Tan, C. Peubey, J.-N. Thpaut, Y. Trmolet, E. V. Hlm, M. Bonavita, L. Isaksen, and M. Fisher, 2016: Era-20c: An atmospheric reanalysis of the twentieth century. *Journal of Climate*, **29**, 4083–4097, doi:10.1175/JCLI-D-15-0556.1.

Pollack, H. N., , P. Y. Shen, and S. Huang, 1996: Inference of ground surface temperature history from subsurface temperature data: interpreting ensembles of borehole logs. *Pageoph*, **147**, 537–550, doi:10.1007/BF00878843.

Pollack, H. N. and S. Huang, 2000: Climate reconstruction from subsurface temperatures. *Annual Review of Earth and Planetary Sciences*, **28**, 339–365, doi:10.1146/annurev.earth.28.1.339.

Pollack, H. N. and J. E. Smerdon, 2004: Borehole climate reconstructions: spatial structure and hemispheric averages. *J. Geophys. Res.*, **109**, D11106, doi:10.1029/2003JD004163.

Pongratz, J., C. Reick, T. Raddatz, and M. Claussen, 2008: A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles*, **22**, GB3018, doi:10.1029/2007GB003153.

Pongratz, J., C. H. Reick, T. Raddatz, and M. Claussen, 2010: Biogeophysical versus biogeochemical climate response to historical anthropogenic land cover change. *Geophysical Research Letters*, **37**, L08702, doi:10.1029/2010GL043010.

Putnam, S. N. and D. S. Chapman, 1996: A geothermal climate change observatory: First year results from emigrant pass in northwest utah. *Journal of Geophysical Research: Solid Earth*, **101**, 21877–21890, doi:10.1029/96JB01903.

Qian, B., E. G. Gregorich, S. Gameda, D. W. Hopkins, and X. L. Wang, 2011: Observed soil temperature trends associated with climate change in canada. *Journal of Geophysical Research: Atmospheres*, **116**, doi:10.1029/2010JD015012.

Ramankutty, N. and J. A. Foley, 1992: Estimating historical chagnes in global land cover: croplands from 1700 to 1992. *Global Biogeochemical Cycles*, **13(4)**, 997–1027.

Reek, T., S. R. Doty, and T. W. Owen, 1992: A deterministic approach to the validation of historical daily temperature and precipitation data from the cooperative network. *Bull. Amer. Meteor. Soc.*, **73**, 753–762, doi:10.1175/1520-0477(1992)073<0753:ADATTV>2.0.CO;2.

Rhoades, D. A. and M. J. Salinger, 1993: Adjustment of temperature and rainfall records for site changes. *Int. J. Climatol.*, **13**, 899–913, doi:10.1002/joc.3370130807.

Río, S. d., L. Herrero, R. Fraile, and A. Penas, 2011: Spatial distribution of recent rainfall trends in spain (1961–2006). *International Journal of Climatology*, **31**, 656–667, doi:10.1002/joc.2111.

Rutherford, S. and M. E. Mann, 2004: Correction to 'optimal surface temperature reconstructions using terrestrial borehole data'. *J. Geophys. Res.*, **109**, D11107, doi: 10.1029/2003JD004290.

Sáez Blázquez, C., A. Farfń Martín, I. Martín Nieto, P. Carrasco García, L. S. Sánchez Pérez, and D. González Aguilera, 2017: Thermal conductivity map of the avila region (spain) based on thermal conductivity measurements of different rock and soil samples. *Geothermics*, **65**, 60 – 71, doi:10.1016/j.geothermics.2016.09.001.

Sancho-Avila, J. M., J. Riesco-Martin, C. Jiménez-Alonso, M. C. Sánchez-de Cos-Escuin, J. Montero-Caldoso, and M. López-Bartolome, 2012: Atlas de radiación solar en españa utilizando datos del saf de clima de eumesat. Technical report, Madrid, Spain.

Santer, B. D., P. W. Thorne, L. Haimberger, K. E. Taylor, T. M. L. Wigley, J. R. Lanzante, S. Solomon, M. Free, P. J. Gleckler, P. D. Jones, T. R. Karl, S. A. Klein, C. Mears, D. Nychka, G. A. Schmidt, S. C. Sherwood, and F. J. Wentz, 2008: Consistency of modelled and observed temperature trends in the tropical troposphere. *International Journal of Climatology*, **28**, 1703–1722, doi:10.1002/joc.1756.

Santer, B. D., T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. Nychka, D. E. Parker, and K. E. Taylor, 2000: Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *Journal of Geophysical Research: Atmospheres*, **105**, 7337–7356, doi:10.1029/1999JD901105.

Schaefer, G. L. and R. F. Paetzold, 2000: SNOTEL (SNOwpack TELemetry) And SCAN (Soil Climate Analysis Network) presented at the Automated Weather Station (AWS) workshop. March 610. Lincoln, Nebr.

Schmidt, G. A., J. H. Jungclaus, C. M. Ammann, E. Bard, P. Braconnot, T. J. Crowley, G. Delaygue, F. Joos, N. A. Krivova, R. Muscheler, B. L. Otto-Bliesner, J. Pongratz, D. T. Shindell, S. K. Solanki, F. Steinhilber, and L. E. A. Vieira, 2011: Climate forcing reconstructions for use in pmip simulations of the last millennium (v1.0). *Geoscientific Model Development*, **4**, 33–45, doi:10.5194/gmd-4-33-2011.

— 2012: Climate forcing reconstructions for use in pmip simulations of the last millennium (v1.1). *Geoscientific Model Development*, **5**, 185–191, doi:10.5194/gmd-5-185-2012.

Schmidt, W. L., W. D. Gosnold, and J. W. Enz, 2001: A decade of air–ground temperature exchange from fargo, north dakota. *Global and Planetary Change*, **29**, 311 – 325, doi:10.1016/S0921-8181(01)00097-2.

Serrano-Notivoli, R., S. Beguería, and M. de Luis, 2019: Stead: a high-resolution daily gridded temperature dataset for spain. *Earth System Science Data*, **11**, 1171–1188, doi:10.5194/essd-11-1171-2019.

Serrano-Notivoli, R., S. Beguería, M. A. Saz, L. A. Longares, and M. de Luis, 2017: Spread: a high-resolution daily gridded precipitation dataset for spain – an extreme events frequency and intensity overview. *Earth System Science Data*, **9**, 721–738, doi:10.5194/essd-9-721-2017.

Shen, P. Y. and A. E. Beck, 1991: Least squares inversion of borehole temperature measurements in functional space. *Journal of Geophysical Research: Solid Earth*, **96**, 19965–19979, doi:10.1029/91JB01883.

Slivinski, L. C., G. P. Compo, J. S. Whitaker, P. D. Sardeshmukh, B. S. Giese, C. McColl, R. Allan, X. Yin, R. Vose, H. Titchner, J. Kennedy, L. J. Spencer, L. Ashcroft, S. Brnnimann, M. Brunet, D. Camuffo, R. Cornes, T. A. Cram, R. Crouthamel, F. Domnguez-Castro, J. E. Freeman, J. Gergis, E. Hawkins, P. D. Jones, S. Jourdain, A. Kaplan, H. Kubota, F. L. Blancq, T.-C. Lee, A. Lorrey, J. Luterbacher, M. Maugeri, C. J. Mock, G. K. Moore, R. Przybylak, C. Pudmenzky, C. Reason, V. C. Slonosky, C. A. Smith, B. Tinz, B. Trewin, M. A. Valente, X. L. Wang, C. Wilkinson, K. Wood, and P. Wyszyski, 2019: Towards a more reliable historical reanalysis: Improvements for version 3 of the twentieth century reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, **145**, 2876–2908, doi:10.1002/qj.3598.

Smerdon, J. E., 2012: Climate models as a test bed for climate reconstruction methods: pseudoproxy experiments. *WIREs Climate Change*, **3**, 63–77, doi:10.1002/wcc.149.

Smerdon, J. E. and H. N. Pollack, 2016: Reconstructing earth's surface temperature over the past 2000 years: the science behind the headlines. *WIREs Climate Change*, **7**, 746–771, doi:https://doi.org/10.1002/wcc.418.

Smerdon, J. E., H. N. Pollack, V. Cermak, J. W. Enz, M. Kresl, J. Safanda, and J. F. Wehmiller, 2004: Air-ground temperature coupling and subsurface propagation of annual temperature signals. *J. Geophys. Res.*, **109**, D21107, doi: 10.1029/2004JD0050506.

— 2006: Daily, seasonal and annual relationships between air and subsurface temperatures. *J. Geophys. Res.*, **111**, D07101, doi: 10.1029/2004JD005578.

Smerdon, J. E., H. N. Pollack, J. W. Enz, and M. J. Lewis, 2003: Conduction-dominated heat transport of the annual temperature signal in soil. *J. Geophys. Res.*, **108**, 2431, doi: 10.1029/2002JB002351.

Smerdon, J. E. and M. Stieglitz, 2006: Simulating heat transport of harmonic temperature signals in the earth's shallow subsurface: Lower-boundary sensitivities. *Geophysical Research Letters*, **33**, doi:10.1029/2006GL026816.

Smith, R., P. Jones, B. Briegleb, F. Bryan, G. Danabasoglu, J. Dennis, J. Dukowicz, C. Eden, B. Fox-Kemper, P. Gent, M. Hecht, S. Jayne, K. L. M. Jochum, W. Large, M. Maltrud, N. Norton, S. Peacock, M. Vertenstein, and S. Yeager, 2010: The Parallel Ocean Program (POP) Reference Manual. Technical Report LAUR-10-01853, Los Alamos National Laboratory, http://www.cesm.ucar.edu/models/cesm1.0/pop2/doc/sci/POPRefManual.pdf (Accessed 2010-02-03).

Solow, A. R., 1987: Testing for climate change: An application of the two-phase regression model. *Journal of Climate and Applied Meteorology*, **26**, 1401–1405, doi:10.1175/1520-0450(1987)026¡1401:TFCCAA¿2.0.CO;2.

Soong, J. L., C. L. Phillips, C. Ledna, C. D. Koven, and M. S. Torn, 2020: Cmip5 models predict rapid and deep soil warming over the 21st century. *Journal of Geophysical Research: Biogeosciences*, **125**, e2019JG005266, doi:10.1029/2019JG005266, e2019JG005266 2019JG005266.

Stieglitz, M., S. J. Dery, V. E. Romanovsky, and T. E. Ostercamp, 2003: The role of snow cover in the warming of the artic permafrost. *Geophys. Res. Lett.*, **30**, 1721, doi:10.1029/2003GL017337.

Tao, W., J. Chen, Z. Li, C. Wang, and C. Zhang, 2012: Impact of aerosols on convective clouds and precipitation. *Reviews of Geophysics*, **50**, doi:10.1029/2011RG000369.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, **106**, 7183–7192, doi:10.1029/2000JD900719.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of cmip5 and the experiment design. *Bulletin of the American Meteorological Society*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Tong, B., Z. Gao, R. Horton, and L. Wang, 2017: Soil apparent thermal diffusivity estimated by conduction and by conduction–convection heat transfer models.

Journal of Hydrometeorology, **18**, 109 – 118, doi:10.1175/JHM-D-16-0086.1.

Trenberth, K. E., 1984: Some effects of finite sample size and persistence on meteorological statistics. part i: autocorrelations. *Mon. Wea. Rev.*, **112**, 2359–2368.

Turcotte, D. and G. Schubert, 2014: *Heat Transfer*, Cambridge University Press. 3 edition, 160229.

Vegas-Cañas, C., J. González-Rouco, J. Navarro-Montesinos, E. García-Bustamante, F. Lucio-Eceiza, E.E. García-Pereira, E. Rodrguez-Camino, Chazarra-Bernabé, and I. A. Álvarez-Arévalo, 2020: An assessment of observed and simulated temperature variability in sierra de guadarrama. *Atmosphere*, **11**, 985, doi:10.3390/atmos11090985.

Vicente-Serrano, S. M., C. Azorin-Molina, A. Sanchez-Lorenzo, E. Morán-Tejeda, J. Lorenzo-Lacruz, J. Revuelto, J. I. López-Moreno, and F. Espejo, 2014: Temporal evolution of surface humidity in spain: recent trends and possible physical mechanisms. *Climate Dynamics*, **42**, 2655–2674, doi:10.1007/s00382-013-1885-7.

Vieira, A., S. K. Solanki, N. A. Krivova, and I. Usoskin, 2011: Evolution of the solar irradiance during the holocene. *Astronomy & Astrophysics*, **531**, A6, doi:10.1051/0004-6361/201015843.

von Randow, C., A. O. Manzi, B. Kruijt, P. J. de Oliveira, F. B. Zanchi, R. L. Silva, M. G. Hodnett, J. H. C. Gash, J. A. Elbers, M. J. Waterloo, F. L. Cardoso, and P. Kabat, 2004: Comparative measurements and seasonal variations in energy and carbon exchange over forest and pasture in south west amazonia. *Theoretical and Applied Climatology*, **78**, 5–26, doi:10.1007/s00704-004-0041-z.

von Storch, H. and F. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 494 pp.

Wang, D., M. Shannon, C. Grieve, and S. Yates, 2000: Soil water and temperature regimes in drip and sprinkler irrigation, and implications to soybean emergence. *Agricultural Water Management*, **43**, 15 – 28, doi:0.1016/S0378-3774(99)00057-8.

Winckler, J., C. H. Reick, S. Luyssaert, A. Cescatti, P. C. Stoy, Q. Lejeune, T. Raddatz, A. Chlond, M. Heidkamp, and J. Pongratz, 2019: Different response of surface temperature and air temperature to deforestation in climate models. *Earth System Dynamics*, **10**, 473–484, doi:10.5194/esd-10-473-2019.

WMO, 2018: Guide to climatological practices. Technical report, World Meteorological Organization, Geneva 2, Switzerland, https://library.wmo.int/doc_num.php?explnum_id=5541 (Accessed 2020-09-17).

Yang, K. and J. Zhang, 2018: Evaluation of reanalysis datasets against observational soil temperature data over china. *Climate Dynamics*, **50**, 317–337,

doi:10.1007/s00382-017-3610-4.

Zhang, H., E. Wang, D. Zhou, Z. Luo, and Z. Zhang, 2016: Rising soil temperature in china and its potential ecological impact. *Scientific Reports*, **16**, 35530, doi:10.1038/srep35530.

Zhang, J., P. Dong, H. Zhang, C. Meng, X. Zhang, J. Hou, and C. Wei, 2019: Low soil temperature reducing the yield of drip irrigated rice in arid area by influencing anther development and pollination. *Journal of Arid Land*, **11**, 419–430, doi:10.1007/s40333-019-0103-9.

Zhang, R. G. G. D., Tingjun. Barry, 2000: Russian Historical Soil Temperature Data. Arctic Data Center. https://data.eol.ucar.edu/dataset/106.ARCSS078 (Accessed 2020-09-12).

Zhao, K. and R. B. Jackson, 2014: Biophysical forcings of land-use changes from potential forestry activities in north america. *Ecological Monographs*, **84**, 329–353.